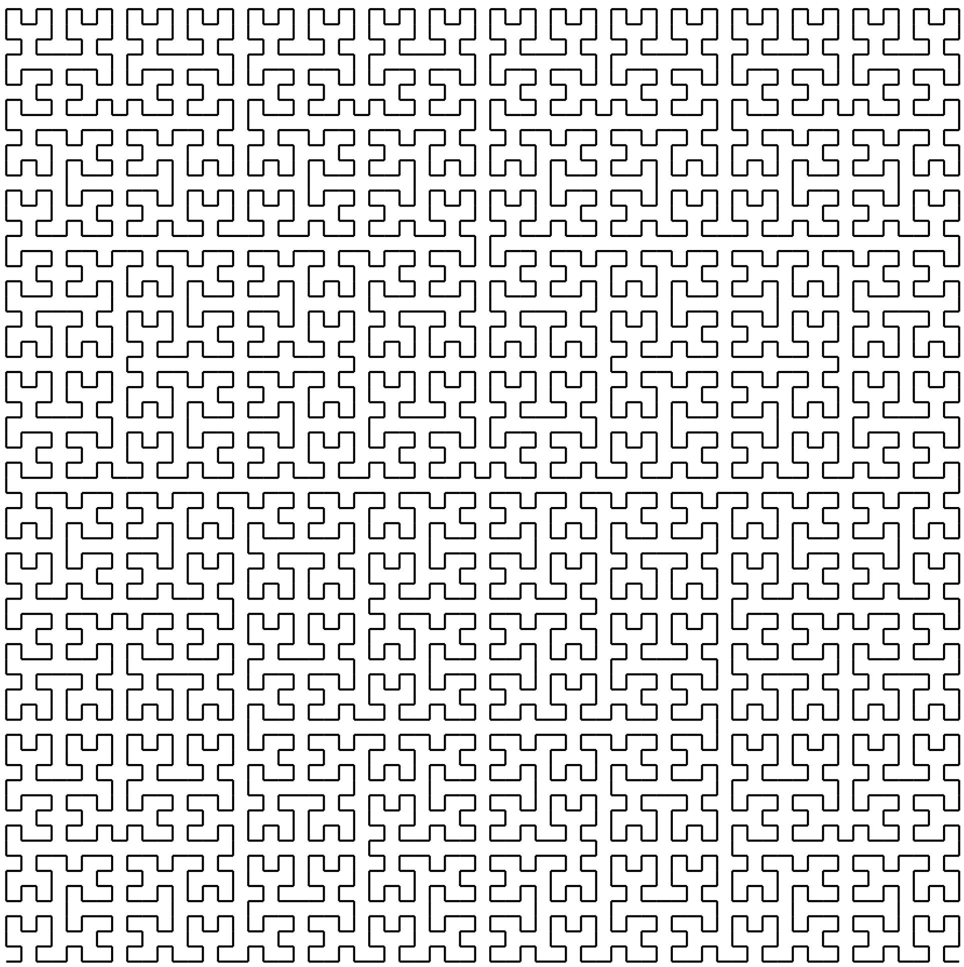
Linear Algebra Notes

David A. SANTOS
dsantos@ccp.edu



May 30, 2008 REVISION

Contents

Preface iv

To the Student v

1 Preliminaries 1
- 1.1 Sets and Notation 1
- 1.2 Partitions and Equivalence Relations 4
- 1.3 Binary Operations 6
- 1.4 $\mathbb{Z}_n$ 9
- 1.5 Fields 13
- 1.6 Functions 15

2 Matrices and Matrix Operations 18
- 2.1 The Algebra of Matrices 18
- 2.2 Matrix Multiplication 22
- 2.3 Trace and Transpose 28
- 2.4 Special Matrices 31
- 2.5 Matrix Inversion 36
- 2.6 Block Matrices 44
- 2.7 Rank of a Matrix 45
- 2.8 Rank and Invertibility 55

3 Linear Equations 65
- 3.1 Definitions 65
- 3.2 Existence of Solutions 70
- 3.3 Examples of Linear Systems 71

4 Vector Spaces 76
- 4.1 Vector Spaces 76
- 4.2 Vector Subspaces 79
- 4.3 Linear Independence 81
- 4.4 Spanning Sets 84
- 4.5 Bases 87
- 4.6 Coordinates 91

5 Linear Transformations 97
- 5.1 Linear Transformations 97
- 5.2 Kernel and Image of a Linear Transformation 99
- 5.3 Matrix Representation 104

6 Determinants 111
- 6.1 Permutations 111
- 6.2 Cycle Notation 114
- 6.3 Determinants 119
- 6.4 Laplace Expansion 129
- 6.5 Determinants and Linear Systems 136

7 Eigenvalues and Eigenvectors 138
- 7.1 Similar Matrices 138
- 7.2 Eigenvalues and Eigenvectors 139
- 7.3 Diagonalisability 143
- 7.4 Theorem of Cayley and Hamilton 147

8 Linear Algebra and Geometry 149
- 8.1 Points and Bi-points in $\mathbb{R}^2$ 149
- 8.2 Vectors in $\mathbb{R}^2$ 152
- 8.3 Dot Product in $\mathbb{R}^2$ 158
- 8.4 Lines on the Plane 164
- 8.5 Vectors in $\mathbb{R}^3$ 169
- 8.6 Planes and Lines in $\mathbb{R}^3$ 174
- 8.7 $\mathbb{R}^n$ 178

A Answers and Hints 182
- Answers and Hints 182

GNU Free Documentation License 262
- 1. APPLICABILITY AND DEFINITIONS 262
- 2. VERBATIM COPYING 262
- 3. COPYING IN QUANTITY 262
- 4. MODIFICATIONS 262
- 5. COMBINING DOCUMENTS 263
- 6. COLLECTIONS OF DOCUMENTS 263
- 7. AGGREGATION WITH INDEPENDENT WORKS 263
- 8. TRANSLATION 263
- 9. TERMINATION 263
- 10. FUTURE REVISIONS OF THIS LICENSE 263

Copyright © 2007 David Anthony SANTOS. Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

Preface

These notes started during the Spring of 2002, when John MAJEWICZ and I each taught a section of Linear Algebra. I would like to thank him for numerous suggestions on the written notes.

The students of my class were: Craig BARIBAULT, Chun CAO, Jacky CHAN, Pho DO, Keith HARMON, Nicholas SELVAGGI, Sanda SHWE, and Huong VU. I must also thank my former student William CARROLL for some comments and for supplying the proofs of a few results.

John's students were David HERNÁNDEZ, Adel JAILILI, Andrew KIM, Jong KIM, Abdelmounaim LAAYOUNI, Aju MATHEW, Nikita MORIN, Thomas NEGRÓN, Latoya ROBINSON, and Saem SOEURN.

Linear Algebra is often a student's first introduction to abstract mathematics. Linear Algebra is well suited for this, as it has a number of beautiful but elementary and easy to prove theorems. My purpose with these notes is to introduce students to the concept of proof in a gentle manner.

David A. Santos dsantos@ccp.edu

To the Student

These notes are provided for your benefit as an attempt to organise the salient points of the course. They are a *very terse* account of the main ideas of the course, and are to be used mostly to refer to central definitions and theorems. The number of examples is minimal, and here you will find few exercises. The *motivation* or informal ideas of looking at a certain topic, the ideas linking a topic with another, the worked-out examples, etc., are given in class. Hence these notes are not a substitute to lectures: **you must always attend to lectures**. The order of the notes may not necessarily be the order followed in the class.

There is a certain algebraic fluency that is necessary for a course at this level. These algebraic prerequisites would be difficult to codify here, as they vary depending on class response and the topic lectured. If at any stage you stumble in Algebra, seek help! I am here to help you!

Tutoring can sometimes help, but bear in mind that whoever tutors you may not be familiar with my conventions. Again, I am here to help! On the same vein, other books may help, but the approach presented here is at times unorthodox and finding alternative sources might be difficult.

Here are more recommendations:

- Read a section before class discussion, in particular, read the definitions.
- Class provides the informal discussion, and you will profit from the comments of your classmates, as well as gain confidence by providing your insights and interpretations of a topic. **Don't be absent!**
- Once the lecture of a particular topic has been given, take a fresh look at the notes of the lecture topic.
- Try to understand a single example well, rather than ill-digest multiple examples.
- Start working on the distributed homework ahead of time.
- **Ask questions during the lecture.** There are two main types of questions that you are likely to ask.
  1. *Questions of Correction: Is that a minus sign there?* If you think that, for example, I have missed out a minus sign or wrote P where it should have been Q,[1] then by all means, ask. No one likes to carry an error till line XLV because the audience failed to point out an error on line I. Don't wait till the end of the class to point out an error. Do it when there is still time to correct it!
  2. *Questions of Understanding: I don't get it!* Admitting that you do not understand something is an act requiring utmost courage. But if you don't, it is likely that many others in the audience also don't. On the same vein, if you feel you can explain a point to an inquiring classmate, I will allow you time in the lecture to do so. The best way to ask a question is something like: "How did you get from the second step to the third step?" or "What does it mean to complete the square?" Asseverations like "I don't understand" do not help me answer your queries. If I consider that you are asking the same questions too many times, it may be that you need extra help, in which case we will settle what to do outside the lecture.
- Don't fall behind! The sequence of topics is closely interrelated, with one topic leading to another.
- The use of calculators is allowed, especially in the occasional lengthy calculations. However, when graphing, you will need to provide algebraic/analytic/geometric support of your arguments. The questions on assignments and exams will be posed in such a way that it will be of no advantage to have a graphing calculator.
- Presentation is critical. Clearly outline your ideas. When writing solutions, outline major steps and write in complete sentences. As a guide, you may try to emulate the style presented in the scant examples furnished in these notes.

[1] My doctoral adviser used to say "I said A, I wrote B, I meant C and it should have been D!

Chapter 1

Preliminaries

1.1 Sets and Notation

1 Definition We will mean by a *set* a collection of well defined members or *elements*.

2 Definition The following sets have special symbols.

$\mathbb{N} = \{0, 1, 2, 3, \ldots\}$	denotes the set of natural numbers.
$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$	denotes the set of integers.
$\mathbb{Q}$	denotes the set of rational numbers.
$\mathbb{R}$	denotes the set of real numbers.
$\mathbb{C}$	denotes the set of complex numbers.
$\varnothing$	denotes the empty set.

3 Definition (Implications) The symbol $\implies$ is read "implies", and the symbol $\iff$ is read "if and only if."

4 Example Prove that between any two rational numbers there is always a rational number.

Solution: ▶ *Let* $(a, c) \in \mathbb{Z}^2$, $(b, d) \in (\mathbb{N} \setminus \{0\})^2$, $\frac{a}{b} < \frac{c}{d}$. *Then* $da < bc$. *Now*

$$ab + ad < ab + bc \implies a(b + d) < b(a + c) \implies \frac{a}{b} < \frac{a + c}{b + d},$$

$$da + dc < cb + cd \implies d(a + c) < c(b + d) \implies \frac{a + c}{b + d} < \frac{c}{d},$$

whence the rational number $\dfrac{a + c}{b + d}$ *lies between* $\frac{a}{b}$ *and* $\frac{c}{d}$. ◀

☞ *We can also argue that the average of two distinct numbers lies between the numbers and so if* $r_1 < r_2$ *are rational numbers, then* $\dfrac{r_1 + r_2}{2}$ *lies between them.*

5 Definition Let A be a set. If a belongs to the set A, then we write $a \in A$, read "a is an element of A." If a does not belong to the set A, we write $a \not\in A$, read "a is not an element of A."

6 Definition (Conjunction, Disjunction, and Negation) The symbol $\vee$ is read "or" (*disjunction*), the symbol $\wedge$ is read "and" (*conjunction*), and the symbol $\neg$ is read "not."

7 Definition (Quantifiers) The symbol $\forall$ is read "for all" (the *universal quantifier*), and the symbol $\exists$ is read "there exists" (the *existential quantifier*).

We have

$$\neg(\forall x \in A, P(x)) \iff (\exists \in A, \neg P(x)) \tag{1.1}$$

$$\neg(\exists \in A, P(x)) \iff (\forall x \in A, \neg P(x)) \tag{1.2}$$

8 Definition (Subset) If $\forall a \in A$ we have $a \in B$, then we write $A \subseteq B$, which we read "A is a subset of B."

In particular, notice that for any set A, $\varnothing \subseteq A$ and $A \subseteq A$. Also

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

$A = B \iff (A \subseteq B) \wedge (B \subseteq A).$

9 Definition The *union* of two sets A and B, is the set

$$A \cup B = \{x : (x \in A) \ \vee \ (x \in B)\}.$$

This is read "A union B." See figure 1.1.

10 Definition The *intersection* of two sets A and B, is

$$A \cap B = \{x : (x \in A) \ \wedge \ (x \in B)\}.$$

This is read "A intersection B." See figure 1.2.

11 Definition The *difference* of two sets A and B, is

$$A \setminus B = \{x : (x \in A) \ \wedge (x \notin B)\}.$$

This is read "A set minus B." See figure 1.3.

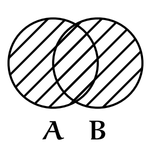


Figure 1.1: $A \cup B$

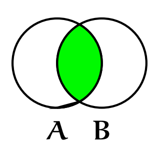


Figure 1.2: $A \cap B$

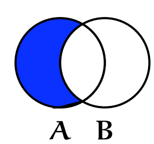


Figure 1.3: $A \setminus B$

12 Example Prove by means of set inclusion that

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C).$$

Solution: ▸ *We have,*

$$\begin{aligned} x \in (A \cup B) \cap C &\iff x \in (A \cup B) \wedge x \in C \\ &\iff (x \in A \vee x \in B) \wedge x \in C \\ &\iff (x \in A \wedge x \in C) \vee (x \in B \wedge x \in C) \\ &\iff (x \in A \cap C) \vee (x \in B \cap C) \\ &\iff x \in (A \cap C) \cup (B \cap C), \end{aligned}$$

which establishes the equality. ◂

13 Definition Let $A_1, A_2, \ldots, A_n$, be sets. The *Cartesian Product* of these n sets is defined and denoted by

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2, \ldots, a_n) : a_k \in A_k\},$$

that is, the set of all ordered n-tuples whose elements belong to the given sets.

☞ *In the particular case when all the A_k are equal to a set A, we write*

$$A_1 \times A_2 \times \cdots \times A_n = A^n.$$

If $a \in A$ and $b \in A$ we write $(a, b) \in A^2$.

14 Definition Let $x \in \mathbb{R}$. The *absolute value of* x—denoted by $|x|$—is defined by

$$|x| = \begin{cases} -x & \text{if } x < 0, \\ x & \text{if } x \geq 0. \end{cases}$$

It follows from the definition that for $x \in \mathbb{R}$,

$$-|x| \leq x \leq |x|. \tag{1.3}$$

$$t \geq 0 \implies |x| \leq t \iff -t \leq x \leq t. \tag{1.4}$$

$$\forall a \in \mathbb{R} \implies \sqrt{a^2} = |a|. \tag{1.5}$$

15 Theorem (Triangle Inequality) Let $(a, b) \in \mathbb{R}^2$. Then

$$|a + b| \leq |a| + |b|. \tag{1.6}$$

Proof: *From 1.3, by addition,*

$$-|a| \leq a \leq |a|$$

to

$$-|b| \leq b \leq |b|$$

we obtain

$$-(|a| + |b|) \leq a + b \leq (|a| + |b|),$$

whence the theorem follows by 1.4. ❑

Homework

Problem 1.1.1 Prove that between any two rational numbers there is an irrational number.

Problem 1.1.2 Prove that $X \setminus (X \setminus A) = X \cap A$.

Problem 1.1.3 Prove that

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

Problem 1.1.4 Prove that

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

Problem 1.1.5 Prove that

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

Problem 1.1.6 Write the union $A \cup B \cup C$ as a *disjoint* union of sets.

Problem 1.1.7 Prove that a set with $n \geq 0$ elements has 2^n subsets.

Problem 1.1.8 Let $(a, b) \in \mathbb{R}^2$. Prove that

$$||a| - |b|| \leq |a - b|.$$

1.2 Partitions and Equivalence Relations

16 Definition Let $S \neq \varnothing$ be a set. A *partition* of S is a collection of non-empty, pairwise disjoint subsets of S whose union is S.

17 Example Let

$$2\mathbb{Z} = \{\ldots, -6, -4, -2, 0, 2, 4, 6, \ldots\} = \overline{0}$$

be the set of even integers and let

$$2\mathbb{Z} + 1 = \{\ldots, -5, -3, -1, 1, 3, 5, \ldots\} = \overline{1}$$

be the set of odd integers. Then

$$(2\mathbb{Z}) \cup (2\mathbb{Z} + 1) = \mathbb{Z}, \quad (2\mathbb{Z}) \cap (2\mathbb{Z} + 1) = \varnothing,$$

and so $\{2\mathbb{Z}, 2\mathbb{Z} + 1\}$ is a partition of $\mathbb{Z}$.

18 Example Let

$$3\mathbb{Z} = \{\ldots - 9, , -6, -3, 0, 3, 6, 9, \ldots\} = \overline{0}$$

be the integral multiples of 3, let

$$3\mathbb{Z} + 1 = \{\ldots, -8, -5, -2, 1, 4, 7, \ldots\} = \overline{1}$$

be the integers leaving remainder 1 upon division by 3, and let

$$3\mathbb{Z} + 2 = \{\ldots, -7, -4, -1, 2, 5, 8, \ldots\} = \overline{2}$$

be integers leaving remainder 2 upon division by 3. Then

$$(3\mathbb{Z}) \cup (3\mathbb{Z} + 1) \cup (3\mathbb{Z} + 2) = \mathbb{Z},$$
$$(3\mathbb{Z}) \cap (3\mathbb{Z} + 1) = \varnothing, \ (3\mathbb{Z}) \cap (3\mathbb{Z} + 2) = \varnothing, (3\mathbb{Z} + 1) \cap (3\mathbb{Z} + 2) = \varnothing,$$

and so $\{3\mathbb{Z}, 3\mathbb{Z} + 1, 3\mathbb{Z} + 2\}$ is a partition of $\mathbb{Z}$.

☞ *Notice that $\overline{0}$ and $\overline{1}$ do not mean the same in examples 17 and 18. Whenever we make use of this notation, the integral divisor must be made explicit.*

19 Example Observe

$$\mathbb{R} = (\mathbb{Q}) \cup (\mathbb{R} \setminus \mathbb{Q}), \quad \varnothing = (\mathbb{Q}) \cap (\mathbb{R} \setminus \mathbb{Q}),$$

which means that the real numbers can be partitioned into the rational and irrational numbers.

20 Definition Let A, B be sets. A *relation* R is a subset of the Cartesian product $A \times B$. We write the fact that $(x, y) \in R$ as $x \sim y$.

21 Definition Let A be a set and R be a relation on $A \times A$. Then R is said to be

- **reflexive** if $(\forall x \in A), x \sim x$,
- **symmetric** if $(\forall (x, y) \in A^2), x \sim y \implies y \sim x$,
- **anti-symmetric** if $(\forall (x, y) \in A^2), (x \sim y) \wedge (y \sim x) \implies x = y$,
- **transitive** if $(\forall (x, y, z) \in A^3), (x \sim y) \wedge (y \sim z) \implies (x \sim z)$.

A relation R which is reflexive, symmetric and transitive is called an *equivalence relation* on A. A relation R which is reflexive, anti-symmetric and transitive is called a *partial order* on A.

22 Example Let S ={All Human Beings}, and define $\sim$ on S as $a \sim b$ if and only if a and b have the same mother. Then $a \sim a$ since any human a has the same mother as himself. Similarly, $a \sim b \implies b \sim a$ and $(a \sim b) \wedge (b \sim c) \implies (a \sim c)$. Therefore $\sim$ is an equivalence relation.

23 Example Let L be the set of all lines on the plane and write $l_1 \sim l_2$ if $l_1 || l_2$ (the line l_1 is parallel to the line l_2). Then $\sim$ is an equivalence relation on L.

24 Example In $\mathbb{Q}$ define the relation $\frac{a}{b} \sim \frac{x}{y} \iff ay = bx$, where we will always assume that the denominators are non-zero. Then $\sim$ is an equivalence relation. For $\frac{a}{b} \sim \frac{a}{b}$ since $ab = ab$. Clearly

$$\frac{a}{b} \sim \frac{x}{y} \implies ay = bx \implies xb = ya \implies \frac{x}{y} \sim \frac{a}{b}.$$

Finally, if $\frac{a}{b} \sim \frac{x}{y}$ and $\frac{x}{y} \sim \frac{s}{t}$ then we have $ay = bx$ and $xt = sy$. Multiplying these two equalities $ayxt = bxsy$. This gives

$$ayxt - bxsy = 0 \implies xy(at - bs) = 0.$$

Now if $x = 0$, we will have $a = s = 0$, in which case trivially $at = bs$. Otherwise we must have $at - bs = 0$ and so $\frac{a}{b} \sim \frac{s}{t}$.

25 Example Let X be a collection of sets. Write $A \sim B$ if $A \subseteq B$. Then $\sim$ is a partial order on X.

26 Example For $(a, b) \in \mathbb{R}^2$ define

$$a \sim b \Leftrightarrow a^2 + b^2 > 2.$$

Determine, with proof, whether $\sim$ is reflexive, symmetric, and/or transitive. Is $\sim$ an equivalence relation?

Solution: ► *Since $0^2 + 0^2 \not> 2$, we have $0 \nsim 0$ and so $\sim$ is not reflexive. Now,*

$$\begin{aligned} a \sim b &\Leftrightarrow a^2 + b^2 \\ &\Leftrightarrow b^2 + a^2 \\ &\Leftrightarrow b \sim a, \end{aligned}$$

so $\sim$ is symmetric. Also $0 \sim 3$ since $0^2 + 3^2 > 2$ and $3 \sim 1$ since $3^2 + 1^2 > 2$. But $0 \nsim 1$ since $0^2 + 1^2 \not> 2$. Thus the relation is not transitive. The relation, therefore, is not an equivalence relation.

◄

27 Definition Let $\sim$ be an equivalence relation on a set S. Then the *equivalence class of* a is defined and denoted by

$$[a] = \{x \in S : x \sim a\}.$$

28 Lemma Let $\sim$ be an equivalence relation on a set S. Then two equivalence classes are either identical or disjoint.

> **Proof:** *We prove that if* $(a, b) \in S^2$, *and* $[a] \cap [b] \neq \varnothing$ *then* $[a] = [b]$. *Suppose that* $x \in [a] \cap [b]$. *Now* $x \in [a] \implies x \sim a \implies a \sim x$, *by symmetry. Similarly,* $x \in [b] \implies x \sim b$. *By transitivity*
>
> $$(a \sim x) \wedge (x \sim b) \implies a \sim b.$$
>
> *Now, if* $y \in [b]$ *then* $b \sim y$. *Again by transitivity,* $a \sim y$. *This means that* $y \in [a]$. *We have shewn that* $y \in [b] \implies y \in [a]$ *and so* $[b] \subseteq [a]$. *In a similar fashion, we may prove that* $[a] \subseteq [b]$. *This establishes the result.* ❑

29 Theorem Let $S \neq \varnothing$ be a set. Any equivalence relation on S induces a partition of S. Conversely, given a partition of S into disjoint, non-empty subsets, we can define an equivalence relation on S whose equivalence classes are precisely these subsets.

> **Proof:** *By Lemma 28, if* $\sim$ *is an equivalence relation on* S *then*
>
> $$S = \bigcup_{a \in S} [a],$$
>
> *and* $[a] \cap [b] = \varnothing$ *if* $a \nsim b$. *This proves the first half of the theorem.*
>
> *Conversely, let*
>
> $$S = \bigcup_{\alpha} S_\alpha, \quad S_\alpha \cap S_\beta = \varnothing \ \textbf{ if } \alpha \neq \beta,$$
>
> *be a partition of* S. *We define the relation* $\approx$ *on* S *by letting* $a \approx b$ *if and only if they belong to the same* S_α. *Since the* S_α *are mutually disjoint, it is clear that* $\approx$ *is an equivalence relation on* S *and that for* $a \in S_\alpha$, *we have* $[a] = S_\alpha$. ❑

Homework

Problem 1.2.1 For $(a, b) \in (\mathbb{Q} \setminus \{0\})^2$ define the relation $\sim$ as follows: $a \sim b \Leftrightarrow \frac{a}{b} \in \mathbb{Z}$. Determine whether this relation is reflexive, symmetric, and/or transitive.

Problem 1.2.2 Give an example of a relation on $\mathbb{Z} \setminus \{0\}$ which is reflexive, but is neither symmetric nor transitive.

Problem 1.2.3 Define the relation $\sim$ in $\mathbb{R}$ by $x \sim y \iff xe^y = ye^x$. Prove that $\sim$ is an equivalence relation.

Problem 1.2.4 Define the relation $\sim$ in $\mathbb{Q}$ by $x \sim y \iff \exists h \in \mathbb{Z}$ such that $x = \dfrac{3y + h}{3}$. [A] Prove that $\sim$ is an equivalence relation. [B] Determine $[x]$, the equivalence of $x \in \mathbb{Q}$. [C] Is $\frac{2}{3} \sim \frac{4}{5}$?

1.3 Binary Operations

30 Definition Let S, T be sets. A *binary operation* is a function

$$\otimes : \begin{array}{ccc} S \times S & \to & T \\ (a, b) & \mapsto & (a, b) \end{array}.$$

We usually use the "infix" notation $a \otimes b$ rather than the "prefix" notation $\otimes(a, b)$. If $S = T$ then we say that the binary operation is *internal* or *closed* and if $S \neq T$ then we say that it is *external*. If

$$a \otimes b = b \otimes a$$

then we say that the operation $\otimes$ is *commutative* and if

$$a \otimes (b \otimes c) = (a \otimes b) \otimes c,$$

we say that it is *associative*. If $\otimes$ is associative, then we can write

$$a \otimes (b \otimes c) = (a \otimes b) \otimes c = a \otimes b \otimes c,$$

without ambiguity.

☞ *We usually omit the sign $\otimes$ and use juxtaposition to indicate the operation $\otimes$. Thus we write ab instead of $a \otimes b$.*

31 Example The operation $+$ (ordinary addition) on the set $\mathbb{Z} \times \mathbb{Z}$ is a commutative and associative closed binary operation.

32 Example The operation $-$ (ordinary subtraction) on the set $\mathbb{N} \times \mathbb{N}$ is a non-commutative, non-associative non-closed binary operation.

33 Example The operation $\otimes$ defined by $a \otimes b = 1 + ab$ on the set $\mathbb{Z} \times \mathbb{Z}$ is a commutative but non-associative internal binary operation. For

$$a \otimes b = 1 + ab = 1 + ba = ba,$$

proving commutativity. Also, $1 \otimes (2 \otimes 3) = 1 \otimes (7) = 8$ and $(1 \otimes 2) \otimes 3 = (3) \otimes 3 = 10$, evincing non-associativity.

34 Definition Let S be a set and $\otimes : S \times S \to S$ be a closed binary operation. The couple $\langle S, \otimes \rangle$ is called an *algebra*.

☞ *When we desire to drop the sign $\otimes$ and indicate the binary operation by juxtaposition, we simply speak of the "algebra S."*

35 Example Both $\langle \mathbb{Z}, + \rangle$ and $\langle \mathbb{Q}, \cdot \rangle$ are algebras. Here $+$ is the standard addition of real numbers and $\cdot$ is the standard multiplication.

36 Example $\langle \mathbb{Z}, - \rangle$ is a non-commutative, non-associative algebra. Here $-$ is the standard subtraction operation on the real numbers

37 Example (Putnam Exam, 1972) Let S be a set and let $*$ be a binary operation of S satisfying the laws $\forall (x, y) \in S^2$

$$x * (x * y) = y, \tag{1.7}$$

$$(y * x) * x = y. \tag{1.8}$$

Shew that $*$ is commutative, but not necessarily associative.

Solution: ▶ *By (1.8)*

$$x * y = ((x * y) * x) * x.$$

By (1.8) again

$$((x * y) * x) * x = ((x * y) * ((x * y) * y)) * x.$$

By (1.7)

$$((x * y) * ((x * y) * y)) * x = (y) * x = y * x,$$

which is what we wanted to prove.

To shew that the operation is not necessarily associative, specialise $S = \mathbb{Z}$ *and* $x * y = -x - y$ *(the opposite of* x *minus* y*). Then clearly in this case* $*$ *is commutative, and satisfies (1.7) and (1.8) but*

$$0 * (0 * 1) = 0 * (-0 - 1) = 0 * (-1) = -0 - (-1) = 1,$$

and

$$(0 * 0) * 1 = (-0 - 0) * 1 = (0) * 1 = -0 - 1 = -1,$$

evincing that the operation is not associative. ◀

38 Definition Let S be an algebra. Then $l \in S$ is called a *left identity* if $\forall s \in S$ we have $ls = s$. Similarly $r \in S$ is called a *right identity* if $\forall s \in S$ we have $sr = s$.

39 Theorem If an algebra S possesses a left identity l and a right identity r then $l = r$.

Proof: *Since* l *is a left identity*

$$r = lr.$$

Since r *is a right identity*

$$l = lr.$$

Combining these two, we gather

$$r = lr = l,$$

whence the theorem follows. ❑

40 Example In $\langle \mathbb{Z}, + \rangle$ the element $0 \in \mathbb{Z}$ acts as an identity, and in $\langle \mathbb{Q}, \cdot \rangle$ the element $1 \in \mathbb{Q}$ acts as an identity.

41 Definition Let S be an algebra. An element $a \in S$ is said to be *left-cancellable* or *left-regular* if $\forall (x, y) \in S^2$

$$ax = ay \implies x = y.$$

Similarly, element $b \in S$ is said to be *right-cancellable* or *right-regular* if $\forall (x, y) \in S^2$

$$xb = yb \implies x = y.$$

Finally, we say an element $c \in S$ is *cancellable* or *regular* if it is both left and right cancellable.

42 Definition Let $\langle S, \otimes \rangle$ and $\langle S, \top \rangle$ be algebras. We say that $\top$ is *left-distributive* with respect to $\otimes$ if

$$\forall (x, y, z) \in S^3, \ x \top (y \otimes z) = (x \top y) \otimes (x \top z).$$

Similarly, we say that $\top$ is *right-distributive* with respect to $\otimes$ if

$$\forall (x, y, z) \in S^3, \ (y \otimes z) \top x = (y \top x) \otimes (z \top x).$$

We say that $\top$ is *distributive* with respect to $\otimes$ if it is both left and right distributive with respect to $\otimes$.

Homework

Problem 1.3.1 Let

$$S = \{x \in \mathbb{Z} : \exists(a, b) \in \mathbb{Z}^2, x = a^3 + b^3 + c^3 - 3abc\}.$$

Prove that S is closed under multiplication, that is, if $x \in S$ and $y \in S$ then $xy \in S$.

Problem 1.3.2 Let $\langle S, \otimes\rangle$ be an associative algebra, let $a \in S$ be a fixed element and define the closed binary operation $\top$ by

$$x \top y = x \otimes a \otimes y.$$

Prove that $\top$ is also associative over $S \times S$.

Problem 1.3.3 On $\mathbb{Q} \cap]-1; 1[$ define the a binary operation $\otimes$

$$a \otimes b = \frac{a + b}{1 + ab},$$

where juxtaposition means ordinary multiplication and $+$ is the ordinary addition of real numbers. Prove that

❶ Prove that $\otimes$ is a closed binary operation on $\mathbb{Q} \cap]-1; 1[$.

❷ Prove that $\otimes$ is both commutative and associative.

❸ Find an element $e \in \mathbb{R}$ such that $(\forall a \in \mathbb{Q} \cap]-1; 1[)\ (e \otimes a = a)$.

❹ Given e as above and an arbitrary element $a \in \mathbb{Q} \cap]-1; 1[$, solve the equation $a \otimes b = e$ for b.

Problem 1.3.4 On $\mathbb{R} \setminus \{1\}$ define the a binary operation $\otimes$

$$a \otimes b = a + b - ab,$$

where juxtaposition means ordinary multiplication and $+$ is the ordinary addition of real numbers. Clearly $\otimes$ is a closed binary operation. Prove that

❶ Prove that $\otimes$ is both commutative and associative.

❷ Find an element $e \in \mathbb{R} \setminus \{1\}$ such that $(\forall a \in \mathbb{R} \setminus \{1\})\ (e \otimes a = a)$.

❸ Given e as above and an arbitrary element $a \in \mathbb{R} \setminus \{1\}$, solve the equation $a \otimes b = e$ for b.

Problem 1.3.5 (Putnam Exam, 1971) Let S be a set and let $\circ$ be a binary operation on S satisfying the two laws

$$(\forall x \in S)(x \circ x = x),$$

and

$$(\forall (x, y, z) \in S^3)((x \circ y) \circ z = (y \circ z) \circ x).$$

Shew that $\circ$ is commutative.

Problem 1.3.6 Define the *symmetric difference* of the sets A, B as $A \triangle B = (A \setminus B) \cup (B \setminus A)$. Prove that $\triangle$ is commutative and associative.

1.4 $\mathbb{Z}_n$

43 Theorem (Division Algorithm) Let $n > 0$ be an integer. Then for any integer a there exist unique integers q (called the *quotient*) and r (called the *remainder*) such that $a = qn + r$ and $0 \leq r < q$.

Proof: *In the proof of this theorem, we use the following property of the integers, called the* well-ordering principle*: any non-empty set of non-negative integers has a smallest element.*

Consider the set

$$S = \{a - bn : b \in \mathbb{Z} \wedge a \geq bn\}.$$

Then S is a collection of nonnegative integers and $S \neq \varnothing$ as $\pm a - 0 \cdot n \in S$ and this is non-negative for one choice of sign. By the Well-Ordering Principle, S has a least element, say r. Now, there must be some $q \in \mathbb{Z}$ such that $r = a - qn$ since $r \in S$. By construction, $r \geq 0$. Let us prove that $r < n$. For assume that $r \geq n$. Then $r > r - n = a - qn - n = a - (q + 1)n \geq 0$, since $r - n \geq 0$. But then $a - (q + 1)n \in S$ and $a - (q + 1)n < r$ which contradicts the fact that r is the smallest member of S. Thus we must have $0 \leq r < n$. To prove that r and q are unique, assume that $q_1 n + r_1 = a = q_2 n + r_2$, $0 \leq r_1 < n$, $0 \leq r_2 < n$. Then $r_2 - r_1 = n(q_1 - q_2)$, that is, n divides $(r_2 - r_1)$. But $|r_2 - r_1| < n$, whence $r_2 = r_1$. From this it also follows that $q_1 = q_2$. This completes the proof. ❑

44 Example If $n = 5$ the Division Algorithm says that we can arrange all the integers in five columns as follows:

$$\begin{array}{ccccc} \vdots & \vdots & \vdots & \vdots & \vdots \\ -10 & -9 & -8 & -7 & -6 \\ -5 & -4 & -3 & -2 & -1 \\ 0 & 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 & 9 \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{array}$$

The arrangement above shews that any integer comes in one of 5 flavours: those leaving remainder 0 upon division by 5, those leaving remainder 1 upon division by 5, etc. We let

$$5\mathbb{Z} = \{\ldots, -15, -10, -5, 0, 5, 10, 15, \ldots\} = \overline{0},$$

$$5\mathbb{Z}+1 = \{\ldots, -14, -9, -4, 1, 6, 11, 16, \ldots\} = \overline{1},$$
$$5\mathbb{Z}+2 = \{\ldots, -13, -8, -3, 2, 7, 12, 17, \ldots\} = \overline{2},$$
$$5\mathbb{Z}+3 = \{\ldots, -12, -7, -2, 3, 8, 13, 18, \ldots\} = \overline{3},$$
$$5\mathbb{Z}+4 = \{\ldots, -11, -6, -1, 4, 9, 14, 19, \ldots\} = \overline{4},$$

and

$$\mathbb{Z}_5 = \{\overline{0}, \overline{1}, \overline{2}, \overline{3}, \overline{4}\}.$$

Let n be a fixed positive integer. Define the relation $\equiv$ by $x \equiv y$ if and only if they leave the same remainder upon division by n. Then clearly $\equiv$ is an equivalence relation. As such it partitions the set of integers $\mathbb{Z}$ into disjoint equivalence classes by Theorem 29. This motivates the following definition.

45 Definition Let n be a positive integer. The n *residue classes* upon division by n are

$$\overline{0} = n\mathbb{Z}, \quad \overline{1} = n\mathbb{Z}+1, \quad \overline{2} = n\mathbb{Z}+2, \quad \ldots, \quad \overline{n-1} = n\mathbb{Z}+n-1.$$

The *set of residue classes modulo* n is

$$\mathbb{Z}_n = \{\overline{0}, \overline{1}, \ldots, \overline{n-1}\}.$$

Our interest is now to define some sort of "addition" and some sort of "multiplication" in $\mathbb{Z}_n$.

46 Theorem (Addition and Multiplication Modulo n) Let n be a positive integer. For $(\overline{a}, \overline{b}) \in (\mathbb{Z}_n)^2$ define $\overline{a} + \overline{b} = \overline{r}$, where r is the remainder of $a+b$ upon division by n. and $\overline{a} \cdot \overline{b} = \overline{t}$, where t is the remainder of ab upon division by n. Then these operations are well defined.

Proof: *We need to prove that given arbitrary representatives of the residue classes, we always obtain the same result from our operations. That is, if* $\overline{a} = \overline{a'}$ *and* $\overline{b} = \overline{b'}$ *then we have* $\overline{a} + \overline{b} = \overline{a'} + \overline{b'}$ *and* $\overline{a} \cdot \overline{b} = \overline{a'} \cdot \overline{b'}$.

Now

$$\overline{a} = \overline{a}' \implies \exists (q, q') \in \mathbb{Z}^2, r \in \mathbb{N}, a = qn + r, \; a' = q'n + r, \; 0 \leq r < n,$$

$$\overline{b} = \overline{b}' \implies \exists (q_1, q_1') \in \mathbb{Z}^2, r_1 \in \mathbb{N}, b = q_1 n + r_1, \; b' = q_1' n + r_1, \; 0 \leq r_1 < n.$$

Hence

$$a+b=(q+q_1)n+r+r_1, \quad a'+b'=(q'+q_1')n+r+r_1,$$

meaning that both $a+b$ *and* $a'+b'$ *leave the same remainder upon division by* n, *and therefore*

$$\overline{a}+\overline{b}=\overline{a+b}=\overline{a'+b'}=\overline{a'}+\overline{b'}.$$

Similarly

$$ab=(qq_1n+qr_1+rq_1)n+rr_1, \quad a'b'=(q'q_1'n+q'r_1+rq_1')n+rr_1,$$

and so both ab *and* $a'b'$ *leave the same remainder upon division by* n, *and therefore*

$$\overline{a}\cdot\overline{b}=\overline{ab}=\overline{a'b'}=\overline{a'}\cdot\overline{b'}.$$

This proves the theorem. ❑

47 Example Let

$$\mathbb{Z}_6=\{\overline{0},\overline{1},\overline{2},\overline{3},\overline{4},\overline{5}\}$$

be the residue classes modulo 6. Construct the natural addition $+$ table for $\mathbb{Z}_6$. Also, construct the natural multiplication $\cdot$ table for $\mathbb{Z}_6$.

Solution: ▶ *The required tables are given in tables 1.1 and 1.2.* ◀

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{5}$	$\overline{5}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$

Table 1.1: Addition table for $\mathbb{Z}_6$.

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{0}$	$\overline{2}$	$\overline{4}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{0}$	$\overline{3}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{2}$	$\overline{0}$	$\overline{4}$	$\overline{2}$
$\overline{5}$	$\overline{0}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Table 1.2: Multiplication table for $\mathbb{Z}_6$.

We notice that even though $\overline{2}\neq\overline{0}$ and $\overline{3}\neq\overline{0}$ we have $\overline{2}\cdot\overline{3}=\overline{0}$ in $\mathbb{Z}_6$. This prompts the following definition.

48 Definition (Zero Divisor) An element $a\neq\overline{0}$ of $\mathbb{Z}_n$ is called a *zero divisor* if $ab=\overline{0}$ for some $b\in\mathbb{Z}_n$.

We will extend the concept of zero divisor later on to various algebras.

49 Example Let

$$\mathbb{Z}_7=\{\overline{0},\overline{1},\overline{2},\overline{3},\overline{4},\overline{5},\overline{6}\}$$

be the residue classes modulo 7. Construct the natural addition $+$ table for $\mathbb{Z}_7$. Also, construct the natural multiplication $\cdot$ table for $\mathbb{Z}_7$

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{5}$	$\overline{5}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{6}$	$\overline{6}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$

Table 1.3: Addition table for $\mathbb{Z}_7$.

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{6}$	$\overline{1}$	$\overline{3}$	$\overline{5}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{6}$	$\overline{2}$	$\overline{5}$	$\overline{1}$	$\overline{4}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{1}$	$\overline{5}$	$\overline{2}$	$\overline{6}$	$\overline{3}$
$\overline{5}$	$\overline{0}$	$\overline{5}$	$\overline{3}$	$\overline{1}$	$\overline{6}$	$\overline{4}$	$\overline{2}$
$\overline{6}$	$\overline{0}$	$\overline{6}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Table 1.4: Multiplication table for $\mathbb{Z}_7$.

Solution: ► *The required tables are given in tables 1.3 and 1.4.* ◄

50 Example Solve the equation

$$\overline{5}x = \overline{3}$$

in $\mathbb{Z}_{11}$.

Solution: ► *Multiplying by $\overline{9}$ on both sides*

$$\overline{45}x = \overline{27},$$

that is,

$$x = \overline{5}.$$

◄

We will use the following result in the next section.

51 Definition Let a, b be integers with one of them different from 0. The greatest common divisor d of a, b, denoted by $d = \mathbf{gcd}(a, b)$ is the largest positive integer that divides both a and b.

52 Theorem (Bachet-Bezout Theorem) The greatest common divisor of any two integers a, b can be written as a linear combination of a and b, i.e., there are integers x, y with

$$\mathbf{gcd}(a, b) = ax + by.$$

Proof: *Let $A = \{ax + by : ax + by > 0, x, y \in \mathbb{Z}\}$. Clearly one of $\pm a, \pm b$ is in A, as one of a, b is not zero. By the Well Ordering Principle, A has a smallest element, say d. Therefore, there are x_0, y_0 such that $d = ax_0 + by_0$. We prove that $d = \mathbf{gcd}(a, b)$. To do this we prove that d divides a and b and that if t divides a and b, then t must also divide then d.*

We first prove that d divides a. By the Division Algorithm, we can find integers $q, r, 0 \leq r < d$ such that $a = dq + r$. Then

$$r = a - dq = a(1 - qx_0) - by_0.$$

If $r > 0$, then $r \in A$ is smaller than the smaller element of A, namely d, a contradiction. Thus $r = 0$. This entails $dq = a$, i.e. d divides a. We can similarly prove that d divides b.

Assume that t divides a and b. Then $a = tm, b = tn$ for integers m, n. Hence $d = ax_0 + bx_0 = t(mx_0 + ny_0)$, that is, t divides d. The theorem is thus proved. ❑

Homework

Problem 1.4.1 Write the addition and multiplication tables of $\mathbb{Z}_{11}$ under natural addition and multiplication modulo 11.

Problem 1.4.2 Solve the equation $\overline{3}x^2 - \overline{5}x + \overline{1} = \overline{0}$ in $\mathbb{Z}_{11}$.

Problem 1.4.3 Solve the equation

$$\overline{5}x^2 = \overline{3}$$

in $\mathbb{Z}_{11}$.

Problem 1.4.4 Prove that if $n > 0$ is a composite integer, $\mathbb{Z}_n$ has zero divisors.

Problem 1.4.5 How many solutions does the equation $x^4 + x^3 + x^2 + x + \overline{1} = \overline{0}$ have in $\mathbb{Z}_{11}$?

1.5 Fields

53 Definition Let $\mathbb{F}$ be a set having at least two elements $0_{\mathbb{F}}$ and $1_{\mathbb{F}}$ ($0_{\mathbb{F}} \neq 1_{\mathbb{F}}$) together with two operations $\cdot$ (multiplication, which we usually represent via juxtaposition) and $+$ (addition). A *field* $\langle \mathbb{F}, \cdot, + \rangle$ is a triplet satisfying the following axioms $\forall (a, b, c) \in \mathbb{F}^3$:

F1 Addition and multiplication are associative:

$$(a + b) + c = a + (b + c), \quad (ab)c = a(bc) \tag{1.9}$$

F2 Addition and multiplication are commutative:

$$a + b = b + a, \quad ab = ba \tag{1.10}$$

F3 The multiplicative operation distributes over addition:

$$a(b + c) = ab + ac \tag{1.11}$$

F4 $0_{\mathbb{F}}$ is the additive identity:

$$0_{\mathbb{F}} + a = a + 0_{\mathbb{F}} = a \tag{1.12}$$

F5 $1_{\mathbb{F}}$ is the multiplicative identity:

$$1_{\mathbb{F}}a = a1_{\mathbb{F}} = a \tag{1.13}$$

F6 Every element has an additive inverse:

$$\exists -a \in \mathbb{F}, \quad a + (-a) = (-a) + a = 0_{\mathbb{F}} \tag{1.14}$$

F7 Every non-zero element has a multiplicative inverse: if $a \neq 0_{\mathbb{F}}$

$$\exists a^{-1} \in \mathbb{F}, \quad aa^{-1} = a^{-1}a = 1_{\mathbb{F}} \tag{1.15}$$

The elements of a field are called *scalars*.

An important property of fields is the following.

54 Theorem A field does not have zero divisors.

Proof: *Assume that* $ab = 0_{\mathbb{F}}$. *If* $a \neq 0_{\mathbb{F}}$ *then it has a multiplicative inverse* a^{-1}. *We deduce*

$$a^{-1}ab = a^{-1}0_{\mathbb{F}} \implies b = 0_{\mathbb{F}}.$$

This means that the only way of obtaining a zero product is if one of the factors is $0_{\mathbb{F}}$. ❑

55 Example $\langle \mathbb{Q}, \cdot, + \rangle$, $\langle \mathbb{R}, \cdot, + \rangle$, and $\langle \mathbb{C}, \cdot, + \rangle$ are all fields. The multiplicative identity in each case is 1 and the additive identity is 0.

56 Example Let

$$\mathbb{Q}(\sqrt{2}) = \{a + \sqrt{2}b : (a, b) \in \mathbb{Q}^2\}$$

and define addition on this set as

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

and multiplication as

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

Then $\langle \mathbb{Q} + \sqrt{2}\mathbb{Q}, , + \rangle$ is a field. Observe $0_{\mathbb{F}} = 0$, $1_{\mathbb{F}} = 1$, that the additive inverse of $a + \sqrt{2}b$ is $-a - \sqrt{2}b$, and the multiplicative inverse of $a + \sqrt{2}b$, $(a, b) \neq (0, 0)$ is

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

Here $a^2 - 2b^2 \neq 0$ since $\sqrt{2}$ is irrational.

57 Theorem If p is a prime, $\langle \mathbb{Z}_p, \cdot, + \rangle$ is a field under $\cdot$ multiplication modulo p and $+$ addition modulo p.

Proof: *Clearly the additive identity is* $\overline{0}$ *and the multiplicative identity is* $\overline{1}$. *The additive inverse of* $\overline{a}$ *is* $\overline{p - a}$. *We must prove that every* $\overline{a} \in \mathbb{Z}_p \setminus \{\overline{0}\}$ *has a multiplicative inverse. Such an* a *satisfies* $\mathbf{gcd}(a, p) = 1$ *and by the Bachet-Bezout Theorem 52, there exist integers* x, y *with* $px + ay = 1$. *In such case we have*

$$\overline{1} = \overline{px + ay} = \overline{ay} = \overline{a} \cdot \overline{y},$$

whence $(\overline{a})^{-1} = \overline{y}$. ❑

58 Definition A field is said to be of *characteristic* $p \neq 0$ if for some positive integer p we have $\forall a \in \mathbb{F}, pa = 0_{\mathbb{F}}$, and no positive integer smaller than p enjoys this property.

If the field does not have characteristic $p \neq 0$ then we say that it is of *characteristic* 0. Clearly $\mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ are of characteristic 0, while $\mathbb{Z}_p$ for prime p, is of characteristic p.

59 Theorem The characteristic of a field is either 0 or a prime.

Proof: *If the characteristic of the field is* 0, *there is nothing to prove. Let* p *be the least positive integer for which* $\forall a \in \mathbb{F}, pa = 0_{\mathbb{F}}$. *Let us prove that* p *must be a prime. Assume that instead we had* $p = st$ *with integers* $s > 1, t > 1$. *Take* $a = 1_{\mathbb{F}}$. *Then we must have* $(st)1_{\mathbb{F}} = 0_{\mathbb{F}}$, *which entails* $(s1_{\mathbb{F}})(t1_{\mathbb{F}}) = 0_{\mathbb{F}}$. *But in a field there are no zero-divisors by Theorem 54, hence either* $s1_{\mathbb{F}} = 0_{\mathbb{F}}$ *or* $t1_{\mathbb{F}} = 0_{\mathbb{F}}$. *But either of these equalities contradicts the minimality of* p. *Hence* p *is a prime.* ❑

Homework

Problem 1.5.1 Consider the set of numbers

$$\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{6}) = \{a + b\sqrt{2} + c\sqrt{3} + d\sqrt{6} : (a, b, c, d) \in \mathbb{Q}^4\}.$$

Assume that $\mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{6})$ is a field under ordinary addition and multiplication. What is the multiplicative inverse of the element $\sqrt{2} + 2\sqrt{3} + 3\sqrt{6}$?

Problem 1.5.2 Let $\mathbb{F}$ be a field and a, b two non-zero elements of $\mathbb{F}$. Prove that

$$-(ab^{-1}) = (-a)b^{-1} = a(-b^{-1}).$$

Problem 1.5.3 Let $\mathbb{F}$ be a field and $a \neq 0_{\mathbb{F}}$. Prove that

$$(-a)^{-1} = -(a^{-1}).$$

Problem 1.5.4 Let $\mathbb{F}$ be a field and a, b two non-zero elements of $\mathbb{F}$. Prove that

$$ab^{-1} = (-a)(-b^{-1}).$$

1.6 Functions

60 Definition By a *function* or a *mapping* from one set to another, we mean a rule or mechanism that assigns to every input element of the first set a unique output element of the second set. We shall call the set of inputs the *domain* of the function, the set of *possible* outputs the *target set* of the function, and the set of *actual* outputs the *image* of the function.

We will generally refer to a function with the following notation:

$$f : \begin{array}{ccc} D & \to & T \\ x & \mapsto & f(x) \end{array}.$$

Here f is the *name of the function*, D is its domain, T is its target set, x is the name of a typical input and $f(x)$ is the output or *image of* x *under* f. We call the assignment $x \mapsto f(x)$ the *assignment rule* of the function. Sometimes x is also called the *independent variable.* The set $f(D) = \{f(a) | a \in D\}$ is called the *image* of f. Observe that $f(D) \subseteq T$.

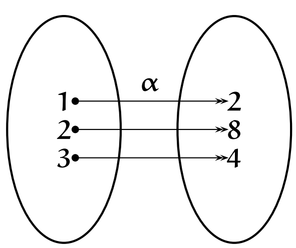


Figure 1.4: An injection.

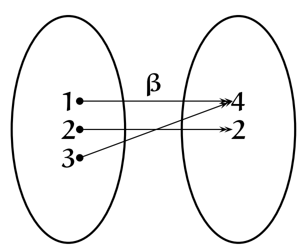


Figure 1.5: Not an injection

61 Definition A function $f : \begin{array}{ccc} X & \to & Y \\ x & \mapsto & f(x) \end{array}$ is said to be *injective* or *one-to-one* if $\forall (a, b) \in X^2$, we have

$$a \neq b \implies f(a) \neq f(b).$$

This is equivalent to saying that
$$f(a) = f(b) \implies a = b.$$

62 Example The function α in the diagram 1.4 is an injective function. The function β represented by the diagram 1.5, however, is not injective, $\beta(3) = \beta(1) = 4$, but $3 \neq 1$.

63 Example Prove that
$$t : \begin{array}{ccc} \mathbb{R}\setminus\{1\} & \to & \mathbb{R}\setminus\{1\} \\ x & \mapsto & \dfrac{x+1}{x-1} \end{array}$$
is an injection.

Solution: ► *Assume* $t(a) = t(b)$. *Then*
$$\begin{array}{rcl} t(a) = t(b) & \implies & \dfrac{a+1}{a-1} = \dfrac{b+1}{b-1} \\ & \implies & (a+1)(b-1) = (b+1)(a-1) \\ & \implies & ab - a + b - 1 = ab - b + a - 1 \\ & \implies & 2a = 2b \\ & \implies & a = b \end{array}$$
We have proved that $t(a) = t(b) \implies a = b$, *which shews that* t *is injective.* ◄

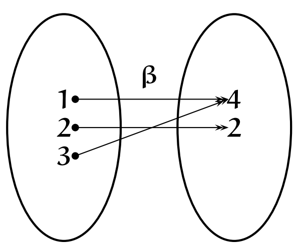


Figure 1.6: A surjection

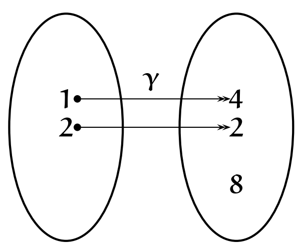


Figure 1.7: Not a surjection

64 Definition A function $f : A \to B$ is said to be *surjective* or *onto* if $(\forall b \in B)\ (\exists a \in A) : f(a) = b$. That is, each element of B has a pre-image in A.

☞ *A function is surjective if its image coincides with its target set. It is easy to see that a graphical criterion for a function to be surjective is that every horizontal line passing through a point of the target set (a subset of the* y*-axis) of the function must also meet the curve.*

65 Example The function β represented by diagram 1.6 is surjective. The function γ represented by diagram 1.7 is not surjective as 8 does not have a preimage.

66 Example Prove that $t : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$ is a surjection.

Solution: ▶ *Since the graph of* t *is that of a cubic polynomial with only one zero, every horizontal line passing through a point in* $\mathbb{R}$ *will eventually meet the graph of* g, *whence* t *is surjective. To prove this analytically, proceed as follows. We must prove that* $(\forall\ b \in \mathbb{R})\ (\exists a)$ *such that* $t(a) = b$. *We choose* a *so that* $a = b^{1/3}$. *Then*

$$t(a) = t(b^{1/3}) = (b^{1/3})^3 = b.$$

Our choice of a *works and hence the function is surjective.* ◀

67 Definition A function is *bijective* if it is both injective and surjective.

Homework

Problem 1.6.1 Prove that

$$h : \begin{array}{ccc} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & x^3 \end{array}$$

is an injection.

Problem 1.6.2 Shew that

$$f : \begin{array}{ccc} \mathbb{R} \setminus \left\{ \dfrac{3}{2} \right\} & \to & \mathbb{R} \setminus \{3\} \\ x & \mapsto & \dfrac{6x}{2x-3} \end{array}$$

is a bijection.

Chapter 2

Matrices and Matrix Operations

2.1 The Algebra of Matrices

68 Definition Let $\langle \mathbb{F}, \cdot, + \rangle$ be a field. An $m \times n$ (m by n) *matrix* A with m rows and n columns with entries over $\mathbb{F}$ is a rectangular array of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

where $\forall (i, j) \in \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}, \quad a_{ij} \in \mathbb{F}$.

☞ *As a shortcut, we often use the notation* $A = [a_{ij}]$ *to denote the matrix* A *with entries* a_{ij}*. Notice that when we refer to the matrix we put parentheses—as in "*$[a_{ij}]$*," and when we refer to a specific entry we do not use the surrounding parentheses—as in "*a_{ij}*."*

69 Example

$$A = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 2 & 3 \end{bmatrix}$$

is a 2×3 matrix and

$$B = \begin{bmatrix} -2 & 1 \\ 1 & 2 \\ 0 & 3 \end{bmatrix}$$

is a 3×2 matrix.

70 Example Write out explicitly the 4×4 matrix $A = [a_{ij}]$ where $a_{ij} = i^2 - j^2$.

Solution: ▶ *This is*

$$A = \begin{bmatrix} 1^2-1^1 & 1^2-2^2 & 1^2-3^2 & 1^2-4^2 \\ 2^2-1^2 & 2^2-2^2 & 2^2-3^2 & 2^2-4^2 \\ 3^2-1^2 & 3^2-2^2 & 3^2-3^2 & 3^2-4^2 \\ 4^2-1^2 & 4^2-2^2 & 4^2-3^2 & 4^2-4^2 \end{bmatrix} = \begin{bmatrix} 0 & -3 & -8 & -15 \\ 3 & 0 & -5 & -12 \\ 8 & 5 & 0 & -7 \\ 15 & 12 & 7 & 0 \end{bmatrix}.$$

◀

71 Definition Let $\langle \mathbb{F}, \cdot, + \rangle$ be a field. We denote by $\mathbf{M}_{m\times n}(\mathbb{F})$ the set of all $m \times n$ matrices with entries over $\mathbb{F}$. $\mathbf{M}_{n\times n}(\mathbb{F})$ is, in particular, the set of all square matrices of size n with entries over $\mathbb{F}$.

72 Definition The $m \times n$ *zero matrix* $\mathbf{0}_{m\times n} \in \mathbf{M}_{m\times n}(\mathbb{F})$ is the matrix with $0_{\mathbb{F}}$'s everywhere,

$$\mathbf{0}_{m\times n} = \begin{bmatrix} 0_{\mathbb{F}} & 0_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \end{bmatrix}.$$

When $m = n$ we write $\mathbf{0}_n$ as a shortcut for $\mathbf{0}_{n\times n}$.

73 Definition The $n \times n$ *identity matrix* $\mathbf{I}_n \in \mathbf{M}_{n\times n}(\mathbb{F})$ is the matrix with $1_{\mathbb{F}}$'s on the main diagonal and $0_{\mathbb{F}}$'s everywhere else,

$$\mathbf{I}_n = \begin{bmatrix} 1_{\mathbb{F}} & 0_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \\ 0_{\mathbb{F}} & 1_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & 1_{\mathbb{F}} & \cdots & 0_{\mathbb{F}} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & 0_{\mathbb{F}} & \cdots & 1_{\mathbb{F}} \end{bmatrix}.$$

74 Definition (Matrix Addition and Multiplication of a Matrix by a Scalar) Let $A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{F})$, $B = [b_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{F})$ and $\alpha \in \mathbb{F}$. The matrix $A + \alpha B$ is the matrix $C \in \mathbf{M}_{m\times n}(\mathbb{F})$ with entries $C = [c_{ij}]$ where $c_{ij} = a_{ij} + \alpha b_{ij}$.

75 Example For $A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \\ 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 1 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}$ we have

$$A + 2B = \begin{bmatrix} -1 & 3 \\ 3 & 3 \\ 0 & 0. \end{bmatrix}.$$

76 Theorem Let $(A, B, C) \in (\mathbf{M}_{m\times n}(\mathbb{F}))^3$ and $(\alpha, \beta) \in \mathbb{F}^2$. Then

M1 $\mathbf{M}_{m\times n}(\mathbb{F})$ is close under matrix addition and scalar multiplication

$$A + B \in \mathbf{M}_{m\times n}(\mathbb{F}), \qquad \alpha A \in \mathbf{M}_{m\times n}(\mathbb{F}) \tag{2.1}$$

M2 Addition of matrices is commutative

$$A + B = B + A \tag{2.2}$$

M3 Addition of matrices is associative

$$A + (B + C) = (A + B) + C \tag{2.3}$$

M4 There is a matrix $\mathbf{0}_{m\times n}$ such that

$$A + \mathbf{0}_{m\times n} \tag{2.4}$$

M5 There is a matrix $-A$ such that

$$A + (-A) = (-A) + A = \mathbf{0}_{m\times n} \tag{2.5}$$

M6 Distributive law

$$\alpha(A + B) = \alpha A + \alpha B \tag{2.6}$$

M7 Distributive law

$$(\alpha + \beta)A = \alpha A + \beta B \tag{2.7}$$

M8

$$1_{\mathbb{F}} A = A \tag{2.8}$$

M9

$$\alpha(\beta A) = (\alpha\beta)A \tag{2.9}$$

Proof: *The theorem follows at once by reducing each statement to an entry-wise and appealing to the field axioms.* ❑

Homework

Problem 2.1.1 Write out explicitly the 3×3 matrix $A = [a_{ij}]$ where $a_{ij} = i^j$.

Problem 2.1.2 Write out explicitly the 3×3 matrix $A = [a_{ij}]$ where $a_{ij} = ij$.

Problem 2.1.3 Let

$$M = \begin{bmatrix} a & -2a & c \\ 0 & -a & b \\ a+b & 0 & -1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 & 2a & c \\ a & b-a & -b \\ a-b & 0 & -1 \end{bmatrix}$$

be square matrices with entries over $\mathbb{R}$. Find $M + N$ and $2M$.

Problem 2.1.4 Determine x and y such that

$$\begin{bmatrix} 3 & x & 1 \\ 1 & 2 & 0 \end{bmatrix} + 2 \begin{bmatrix} 2 & 1 & 3 \\ 5 & x & 4 \end{bmatrix} = \begin{bmatrix} 7 & 3 & 7 \\ 11 & y & 8 \end{bmatrix}.$$

Problem 2.1.5 Determine 2×2 matrices A and B such that

$$2A - 5B = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}, \quad -2A + 6B = \begin{bmatrix} 4 & 2 \\ 6 & 0 \end{bmatrix}.$$

Problem 2.1.6 Let $A = [a_{ij}] \in \mathbf{M}_{n \times n}(\mathbb{R})$. Prove that

$$\min_j \max_i a_{ij} \geq \max_i \min_j a_{ij}.$$

Problem 2.1.7 A person goes along the rows of a movie theater and asks the tallest person of each row to stand up. Then he selects the shortest of these people, who we will call the *shortest giant*. Another person goes along the rows and asks the shortest person to stand up and from these he selects the tallest, which we will call the *tallest midget*. Who is taller, the tallest midget or the shortest giant?

Problem 2.1.8 (Putnam Exam, 1959) Choose five elements from the matrix

$$\begin{bmatrix} 11 & 17 & 25 & 19 & 16 \\ 24 & 10 & 13 & 15 & 3 \\ 12 & 5 & 14 & 2 & 18 \\ 23 & 4 & 1 & 8 & 22 \\ 6 & 20 & 7 & 21 & 9 \end{bmatrix},$$

no two coming from the same row or column, so that the minimum of these five elements is as large as possible.

2.2 Matrix Multiplication

77 Definition Let $\mathbf{A} = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{F})$ and $\mathbf{B} = [b_{ij}] \in \mathbf{M}_{n\times p}(\mathbb{F})$. Then the matrix product $\mathbf{AB}$ is defined as the matrix $\mathbf{C} = [c_{ij}] \in \mathbf{M}_{m\times p}(\mathbb{F})$ with entries $c_{ij} = \sum_{l=1}^{n} a_{il}b_{lj}$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \cdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1p} \\ b_{21} & \cdots & b_{2j} & \cdots & b_{2p} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} c_{11} & \cdots & c_{1p} \\ c_{21} & \cdots & c_{2p} \\ \vdots & \cdots & \vdots \\ \cdots & c_{ij} & \cdots \\ \vdots & \cdots & \vdots \\ c_{m1} & \cdots & c_{mp} \end{bmatrix}.$$

☞ *Observe that we use juxtaposition rather than a special symbol to denote matrix multiplication. This will simplify notation.In order to obtain the ij-th entry of the matrix $\mathbf{AB}$ we multiply elementwise the i-th row of $\mathbf{A}$ by the j-th column of $\mathbf{B}$. Observe that $\mathbf{AB}$ is a $m \times p$ matrix.*

78 Example Let $M = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $N = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ be matrices over $\mathbb{R}$. Then

$$MN = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1\cdot 5 + 2\cdot 7 & 1\cdot 6 + 2\cdot 8 \\ 3\cdot 5 + 4\cdot 7 & 3\cdot 6 + 4\cdot 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix},$$

and

$$NM = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 5\cdot 1 + 6\cdot 3 & 5\cdot 2 + 6\cdot 4 \\ 7\cdot 1 + 8\cdot 3 & 7\cdot 2 + 8\cdot 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}.$$

Hence, in particular, matrix multiplication is not necessarily commutative.

79 Example We have

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

over $\mathbb{R}$. Observe then that the product of two non-zero matrices may be the zero matrix.

80 Example Consider the matrix

$$A = \begin{bmatrix} \overline{2} & \overline{1} & \overline{3} \\ \overline{0} & \overline{1} & \overline{1} \\ \overline{4} & \overline{4} & \overline{0} \end{bmatrix}$$

with entries over $\mathbb{Z}_5$. Then

$$A^2 = \begin{bmatrix} \overline{2} & \overline{1} & \overline{3} \\ \overline{0} & \overline{1} & \overline{1} \\ \overline{4} & \overline{4} & \overline{0} \end{bmatrix} \begin{bmatrix} \overline{2} & \overline{1} & \overline{3} \\ \overline{0} & \overline{1} & \overline{1} \\ \overline{4} & \overline{4} & \overline{0} \end{bmatrix} = \begin{bmatrix} \overline{1} & \overline{0} & \overline{2} \\ \overline{4} & \overline{0} & \overline{1} \\ \overline{3} & \overline{3} & \overline{1} \end{bmatrix}.$$

☞ *Even though matrix multiplication is not necessarily commutative, it is associative.*

81 Theorem If $(A, B, C) \in \mathbf{M}_{m\times n}(\mathbb{F}) \times \mathbf{M}_{n\times r}(\mathbb{F}) \times \mathbf{M}_{r\times s}(\mathbb{F})$ we have

$$(AB)C = A(BC),$$

i.e., matrix multiplication is associative.

Proof: *To shew this we only need to consider the* ij*-th entry of each side, appeal to the associativity of the underlying field* $\mathbb{F}$ *and verify that both sides are indeed equal to*

$$\sum_{k=1}^{n} \sum_{k'=1}^{r} a_{ik} b_{kk'} c_{k'j}.$$

❑

☞ *By virtue of associativity, a square matrix commutes with its powers, that is, if* $A \in \mathbf{M}_{n\times n}(\mathbb{F})$*, and* $(r, s) \in \mathbb{N}^2$*, then* $(A^r)(A^s) = (A^s)(A^r) = A^{r+s}$.

82 Example Let $A \in \mathbf{M}_{3\times 3}(\mathbb{R})$ be given by

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Demonstrate, using induction, that $A^n = 3^{n-1}A$ for $n \in \mathbb{N}, n \geq 1$.

Solution: ▶ *The assertion is trivial for* $n = 1$. *Assume its truth for* $n - 1$, *that is, assume* $A^{n-1} = 3^{n-2}A$. *Observe that*

$$A^2 = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 3 & 3 \\ 3 & 3 & 3 \end{bmatrix} = 3A.$$

Now

$$A^n = AA^{n-1} = A(3^{n-2}A) = 3^{n-2}A^2 = 3^{n-2}3A = 3^{n-1}A,$$

and so the assertion is proved by induction. ◀

83 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then there is a unique identity matrix. That is, if $E \in \mathbf{M}_{n\times n}(\mathbb{F})$ is such that $AE = EA = A$, then $E = \mathbf{I}_n$.

Proof: *It is clear that for any* $A \in \mathbf{M}_{n\times n}(\mathbb{F})$, $A\mathbf{I}_n = \mathbf{I}_nA = A$. *Now because* E *is an identity,* $E\mathbf{I}_n = \mathbf{I}_n$. *Because* $\mathbf{I}_n$ *is an identity,* $E\mathbf{I}_n = E$. *Whence*

$$\mathbf{I}_n = E\mathbf{I}_n = E,$$

demonstrating uniqueness. ❑

84 Example Let $A = [a_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{R})$ be such that $a_{ij} = 0$ for $i > j$ and $a_{ij} = 1$ if $i \leq j$. Find A^2.

Solution: ▶ *Let* $A^2 = B = [b_{ij}]$. *Then*

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

Observe that the i*-th row of* A *has* $i-1$ *0's followed by* $n-i+1$ *1's, and the* j*-th column of* A *has* j *1's followed by* $n-j$ *0's. Therefore if* $i-1 > j$, *then* $b_{ij} = 0$. *If* $i \leq j+1$, *then*

$$b_{ij} = \sum_{k=i}^{j} a_{ik}a_{kj} = j-i+1.$$

This means that

$$A^2 = \begin{bmatrix} 1 & 2 & 3 & 4 & \cdots & n-1 & n \\ 0 & 1 & 2 & 3 & \cdots & n-2 & n-1 \\ 0 & 0 & 1 & 2 & \cdots & n-3 & n-2 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

◀

Homework

Problem 2.2.1 Determine the product

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

.

Problem 2.2.2 Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$, $B = \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix}$. Find AB and BA.

Problem 2.2.3 Find $a+b+c$ if $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} a & a & a \\ b & b & b \\ c & c & c \end{bmatrix}$.

Problem 2.2.4 Let $N = \begin{bmatrix} 0 & -2 & -3 & -4 \\ 0 & 0 & -2 & -3 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Find N^{2008}.

Problem 2.2.5 Let

$$A = \begin{bmatrix} \overline{2} & \overline{3} & \overline{4} & \overline{1} \\ \overline{1} & \overline{2} & \overline{3} & \overline{4} \\ \overline{4} & \overline{1} & \overline{2} & \overline{3} \\ \overline{3} & \overline{4} & \overline{1} & \overline{2} \end{bmatrix}, \quad B = \begin{bmatrix} \overline{1} & \overline{1} & \overline{1} & \overline{1} \\ \overline{1} & \overline{1} & \overline{1} & \overline{1} \\ \overline{1} & \overline{1} & \overline{1} & \overline{1} \\ \overline{1} & \overline{1} & \overline{1} & \overline{1} \end{bmatrix}$$

be matrices in $\mathbf{M}_{4\times 4}(\mathbb{Z}_5)$. Find the products AB and BA.

Problem 2.2.6 Let x be a real number, and put

$$m(x) = \begin{bmatrix} 1 & 0 & x \\ -x & 1 & -\frac{x^2}{2} \\ 0 & 0 & 1 \end{bmatrix}.$$

If a, b are real numbers, prove that

1. $m(a)m(b) = m(a+b)$.
2. $m(a)m(-a) = \mathbf{I}_3$, the 3×3 identity matrix.

Problem 2.2.7 A square matrix X is called *idempotent* if $X^2 = X$. Prove that if $AB = A$ and $BA = B$ then A and B are idempotent.

Problem 2.2.8 Let

$$A = \begin{bmatrix} 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Calculate the value of the infinite series

$$\mathbf{I}_3 + A + A^2 + A^3 + \cdots.$$

Problem 2.2.9 Solve the equation

$$\begin{bmatrix} -4 & x \\ -x & 4 \end{bmatrix}^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

over $\mathbb{R}$.

Problem 2.2.10 Prove or disprove! If $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$ are such that $AB = \mathbf{0}_n$, then also $BA = \mathbf{0}_n$.

Problem 2.2.11 Prove or disprove! For all matrices $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$,

$$(A + B)(A - B) = A^2 - B^2.$$

Problem 2.2.12 Consider the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & x \end{bmatrix}$, where x is a real number. Find the value of x such that there are non-zero 2×2 matrices B such that $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.

Problem 2.2.13 Prove, using mathematical induction, that $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$.

Problem 2.2.14 Let $M = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. Find M^6.

Problem 2.2.15 Let $A = \begin{bmatrix} 0 & 3 \\ 2 & 0 \end{bmatrix}$. Find, with proof, A^{2003}.

Problem 2.2.16 Let $(A, B, C) \in \mathbf{M}_{l\times m}(\mathbb{F}) \times \mathbf{M}_{m\times n}(\mathbb{F}) \times \mathbf{M}_{m\times n}(\mathbb{F})$ and $\alpha \in \mathbb{F}$. Prove that

$$A(B + C) = AB + AC,$$

$$(A + B)C = AC + BC,$$

$$\alpha(AB) = (\alpha A)B = A(\alpha B).$$

Problem 2.2.17 Let $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$ be given by

$$A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}.$$

Demonstrate, using induction, that for $n \in \mathbb{N}, n \geq 1$.

$$A^n = \begin{bmatrix} \cos n\alpha & -\sin n\alpha \\ \sin n\alpha & \cos n\alpha \end{bmatrix}.$$

Problem 2.2.18 A matrix $A = [a_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{R})$ is said to be *checkered* if $a_{ij} = 0$ when $(j - i)$ is odd. Prove that the sum and the product of two checkered matrices is checkered.

Problem 2.2.19 Let $A \in \mathbf{M}_{3\times 3}(\mathbb{R})$,

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Prove that

$$A^n = \begin{bmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}.$$

Problem 2.2.20 Let $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$ and k be a positive integer such that $A^k = \mathbf{0}_n$. If $AB = B$ prove that $B = \mathbf{0}_n$.

Problem 2.2.21 Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Demonstrate that

$$A^2 - (a + d)A + (ad - bc)\mathbf{I}_2 = \mathbf{0}_2$$

.

Problem 2.2.22 Let $A \in \mathbf{M}_2(\mathbb{F})$ and let $k \in \mathbb{Z}, k > 2$. Prove that $A^k = \mathbf{0}_2$ if and only if $A^2 = \mathbf{0}_2$.

Problem 2.2.23 Find all matrices $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$ such that $A^2 = \mathbf{0}_2$

Problem 2.2.24 Find all matrices $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$ such that $A^2 = \mathbf{I}_2$

Problem 2.2.25 Find a solution $X \in \mathbf{M}_{2\times 2}(\mathbb{R})$ for

$$X^2 - 2X = \begin{bmatrix} -1 & 0 \\ 6 & 3 \end{bmatrix}.$$

Problem 2.2.26 Find, with proof, a 4×4 **non-zero** matrix A such that

$$A \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Problem 2.2.27 Let X be a 2×2 matrices with real number entries. Solve the equation

$$X^2 + X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Problem 2.2.28 Prove, by means of induction that for the following $n \times n$ we have

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}^3 = \begin{bmatrix} 1 & 3 & 6 & \cdots & \frac{n(n+1)}{2} \\ 0 & 1 & 3 & \cdots & \frac{(n-1)n}{2} \\ 0 & 0 & 1 & \cdots & \frac{(n-2)(n-1)}{2} \\ \vdots & \vdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Problem 2.2.29 Let

$$A = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{bmatrix}.$$

Conjecture a formula for A^n and prove it using induction.

2.3 Trace and Transpose

85 Definition Let $A = [a_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then the *trace* of A, denoted by $\mathbf{tr}(A)$ is the sum of the diagonal elements of A, that is

$$\mathbf{tr}(A) = \sum_{k=1}^{n} a_{kk}.$$

86 Theorem Let $A = [a_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{F}), B = [b_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then

$$\mathbf{tr}(A + B) = \mathbf{tr}(A) + \mathbf{tr}(B), \tag{2.10}$$

$$\mathbf{tr}(AB) = \mathbf{tr}(BA). \tag{2.11}$$

Proof: *The first assertion is trivial. To prove the second, observe that* $AB = (\sum_{k=1}^{n} a_{ik}b_{kj})$ *and* $BA = (\sum_{k=1}^{n} b_{ik}a_{kj})$. *Then*

$$\mathbf{tr}(AB) = \sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}b_{ki} = \sum_{k=1}^{n}\sum_{i=1}^{n} b_{ki}a_{ik} = \mathbf{tr}(BA),$$

whence the theorem follows. ❑

87 Example Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$. Shew that A can be written as the sum of two matrices whose trace is different from 0.

Solution: ► *Write*

$$A = (A - \alpha \mathbf{I}_n) + \alpha \mathbf{I}_n.$$

Now, $\mathbf{tr}(A - \alpha\mathbf{I}_n) = \mathbf{tr}(A) - n\alpha$ *and* $\mathbf{tr}(\alpha\mathbf{I}_n) = n\alpha$. *Thus it suffices to take* $\alpha \neq \dfrac{\mathbf{tr}(A)}{n}, \alpha \neq 0$. *Since* $\mathbb{R}$ *has infinitely many elements, we can find such an* α.

◄

88 Example Let A, B be square matrices of the same size and over the same field of characteristic 0. Is it possible that $AB - BA = \mathbf{I}_n$? Prove or disprove!

Solution: ► *This is impossible. For if, taking traces on both sides*

$$0 = \mathbf{tr}(AB) - \mathbf{tr}(BA) = \mathbf{tr}(AB - BA) = \mathbf{tr}(\mathbf{I}_n) = n$$

a contradiction, since $n > 0$. ◄

89 Definition The *transpose* of a matrix of a matrix $A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{F})$ is the matrix $A^T = B = [b_{ij}] \in \mathbf{M}_{n\times m}(\mathbb{F})$, where $b_{ij} = a_{ji}$.

90 Example If

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix},$$

with entries in $\mathbb{R}$, then

$$M^T = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}.$$

91 Theorem Let

$$A = [a_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{F}),\ B = [b_{ij}] \in \mathbf{M}_{m\times n}(\mathbb{F}),\ C = [c_{ij}] \in \mathbf{M}_{n\times r}(\mathbb{F}),\ \alpha \in \mathbb{F}, u \in \mathbb{N}.$$

Then

$$A^{TT} = A, \tag{2.12}$$

$$(A + \alpha B)^T = A^T + \alpha B^T, \tag{2.13}$$

$$(AC)^T = C^T A^T, \tag{2.14}$$

$$(A^u)^T = (A^T)^u. \tag{2.15}$$

Proof: *The first two assertions are obvious, and the fourth follows from the third by using induction. To prove the third put* $A^T = (\alpha_{ij})$, $\alpha_{ij} = a_{ji}$, $C^T = (\gamma_{ij})$, $\gamma_{ij} = c_{ji}$, $AC = (u_{ij})$ *and* $C^T A^T = (v_{ij})$. *Then*

$$u_{ij} = \sum_{k=1}^{n} a_{ik}c_{kj} = \sum_{k=1}^{n} \alpha_{ki}\gamma_{jk} = \sum_{k=1}^{n} \gamma_{jk}\alpha_{ki} = v_{ji},$$

whence the theorem follows. ❑

92 Definition A square matrix $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ is *symmetric* if $A^T = A$. A matrix $B \in \mathbf{M}_{n\times n}(\mathbb{F})$ is *skew-symmetric* if $B^T = -B$.

93 Example Let A, B be square matrices of the same size, with A symmetric and B skew-symmetric. Prove that the matrix A^2BA^2 is skew-symmetric.

Solution: ► *We have*

$$(A^2BA^2)^T = (A^2)^T(B)^T(A^2)^T = A^2(-B)A^2 = -A^2BA^2.$$

◄

94 Theorem Let $\mathbb{F}$ be a field of characteristic different from 2. Then any square matrix A can be written as the sum of a symmetric and a skew-symmetric matrix.

Proof: *Observe that*
$$(A + A^T)^T = A^T + A^{TT} = A^T + A,$$
and so $A + A^T$ *is symmetric. Also,*
$$(A - A^T)^T = A^T - A^{TT} = -(A - A^T),$$
and so $A - A^T$ *is skew-symmetric. We only need to write* A *as*
$$A = (2^{-1})(A + A^T) + (2^{-1})(A - A^T)$$
to prove the assertion. ❑

95 Example Find, with proof, a square matrix A with entries in $\mathbb{Z}_2$ such A is not the sum of a symmetric and and anti-symmetric matrix.

Solution: ► *In* $\mathbb{Z}_2$ *every symmetric matrix is also anti-symmetric, since* $-x = x$*. Thus it is enough to take a non-symmetric matrix, for example, take* $A = \begin{bmatrix} \overline{0} & \overline{1} \\ \overline{0} & \overline{0} \end{bmatrix}$. ◄

Homework

Problem 2.3.1 Write
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \in \mathbf{M}_{3\times 3}(\mathbb{R})$$
as the sum of two 3×3 matrices $\mathbf{E}_1, \mathbf{E}_2$, with $\mathbf{tr}(\mathbf{E}_2) = 10$.

Problem 2.3.2 Give an example of two matrices $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$ and $B \in \mathbf{M}_{2\times 2}(\mathbb{R})$ that *simultaneously* satisfy the following properties:

1. $A \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $B \neq \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
2. $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ and $BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$.
3. $\mathbf{tr}(A) = \mathbf{tr}(B) = 2$.
4. $A = A^T$ and $B = B^T$.

Problem 2.3.3 Shew that there are no matrices $(A, B, C, D) \in (\mathbf{M}_{n\times n}(\mathbb{R}))^4$ such that
$$AC + DB = \mathbf{I}_n,$$
$$CA + BD = \mathbf{0}_n.$$

Problem 2.3.4 Let $(A, B) \in (\mathbf{M}_{2\times 2}(\mathbb{R}))^2$ be symmetric matrices. Must their product AB be symmetric? Prove or disprove!

Problem 2.3.5 Given square matrices $(A, B) \in (\mathbf{M}_{7\times 7}(\mathbb{R}))^2$ such that $\mathbf{tr}\left(A^2\right) = \mathbf{tr}\left(B^2\right) = 1$, and
$$(A - B)^2 = 3\mathbf{I}_7,$$
find $\mathbf{tr}(BA)$.

Problem 2.3.6 Consider the matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{M}_{2\times 2}(\mathbb{R})$. Find necessary and sufficient conditions on a, b, c, d so that $\mathbf{tr}\left(A^2\right) = (\mathbf{tr}(A))^2$.

Problem 2.3.7 Given a square matrix $A \in \mathbf{M}_{4\times 4}(\mathbb{R})$ such that $\mathbf{tr}\left(A^2\right) = -4$, and
$$(A - \mathbf{I}_4)^2 = 3\mathbf{I}_4,$$
find $\mathbf{tr}(A)$.

Problem 2.3.8 Prove or disprove! If A, B are square matrices of the same size, then it is always true that $\mathbf{tr}(AB) = \mathbf{tr}(A)\,\mathbf{tr}(B)$.

Problem 2.3.9 Prove or disprove! If $(A, B, C) \in (\mathbf{M}_{3\times 3}(\mathbb{F}))^3$ then $\mathbf{tr}(ABC) = \mathbf{tr}(BAC)$.

Problem 2.3.10 Let A be a square matrix. Prove that the matrix AA^T is symmetric.

Problem 2.3.11 Let A, B be square matrices of the same size, with A symmetric and B skew-symmetric. Prove that the matrix $AB - BA$ is symmetric.

Problem 2.3.12 Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$. Prove that $\mathbf{tr}\left(AA^T\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2$.

Problem 2.3.13 Let $X \in \mathbf{M}_{n\times n}(\mathbb{R})$. Prove that if $XX^T = \mathbf{0}_n$ then $X = \mathbf{0}_n$.

Problem 2.3.14 Let m, n, p be positive integers and $A \in \mathbf{M}_{m\times n}(\mathbb{R})$, $B \in \mathbf{M}_{n\times p}(\mathbb{R})$, $C \in \mathbf{M}_{p\times m}(\mathbb{R})$. Prove that $(BA)^T A = (CA)^T A \implies BA = CA$.

2.4 Special Matrices

96 Definition The *main diagonal* of a square matrix $A = [a_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{F})$ is the set $\{a_{ii} : i \leq n\}$. The *counter diagonal* of a square matrix $A = [a_{ij}] \in \mathbf{M}_{n\times n}(\mathbb{F})$ is the set $\{a_{(n-i+1)i} : i \leq n\}$.

97 Example The main diagonal of the matrix

$$A = \begin{bmatrix} 0 & 1 & 5 \\ 3 & 2 & 4 \\ 9 & 8 & 7 \end{bmatrix}$$

is the set $\{0, 2, 7\}$. The counter diagonal of A is the set $\{5, 2, 9\}$.

98 Definition A square matrix is a *diagonal* matrix if every entry off its main diagonal is $0_{\mathbb{F}}$.

99 Example The matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

is a diagonal matrix.

100 Definition A square matrix is a *scalar* matrix if it is of the form $\alpha \mathbf{I}_n$ for some scalar α.

101 Example The matrix

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} = 4\mathbf{I}_3$$

is a scalar matrix.

102 Definition $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ is said to be *upper triangular* if

$$(\forall (i, j) \in \{1, 2, \cdots, n\}^2), (i > j,\ a_{ij} = 0_{\mathbb{F}}),$$

that is, every element below the main diagonal is $0_{\mathbb{F}}$. Similarly, A is *lower triangular* if

$$(\forall (i, j) \in \{1, 2, \cdots, n\}^2), (i < j,\ a_{ij} = 0_{\mathbb{F}}),$$

that is, every element above the main diagonal is $0_{\mathbb{F}}$.

103 Example The matrix $A \in \mathbf{M}_{3\times 4}(\mathbb{R})$ shewn is upper triangular and $B \in \mathbf{M}_{4\times 4}(\mathbb{R})$ is lower triangular.

$$A = \begin{bmatrix} 1 & a & b & c \\ 0 & 2 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & a & 0 & 0 \\ 0 & 2 & 3 & 0 \\ 1 & 1 & t & 1 \end{bmatrix}$$

104 Definition The *Kronecker delta* δ_{ij} is defined by

$$\delta_{ij} = \begin{cases} 1_{\mathbb{F}} & \textbf{if } i = j \\ 0_{\mathbb{F}} & \textbf{if } i \neq j \end{cases}$$

105 Definition The set of matrices $\mathbf{E}_{ij} \in \mathbf{M}_{m\times n}(\mathbb{F}), \mathbf{E}_{ij} = (e_{rs})$ such that $e_{ij} = 1_{\mathbb{F}}$ and $e_{i'j'} = 0_{\mathbb{F}}, (i', j') \neq (i, j)$ is called the set of *elementary matrices*. Observe that in fact $e_{rs} = \delta_{ir}\delta_{sj}$.

Elementary matrices have interesting effects when we pre-multiply and post-multiply a matrix by them.

106 Example Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}, \quad \mathbf{E}_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then

$$\mathbf{E}_{23}A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 9 & 10 & 11 & 12 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad AE_{23} = \begin{bmatrix} 0 & 0 & 2 \\ 0 & 0 & 6 \\ 0 & 0 & 10 \end{bmatrix}.$$

107 Theorem (Multiplication by Elementary Matrices) Let $\mathbf{E}_{ij} \in \mathbf{M}_{m\times n}(\mathbb{F})$ be an elementary matrix, and let $A \in \mathbf{M}_{n\times m}(\mathbb{F})$. Then $\mathbf{E}_{ij}A$ has as its i-th row the j-th row of A and $0_{\mathbb{F}}$'s everywhere else. Similarly, $A\mathbf{E}_{ij}$ has as its j-th column the i-th column of A and $0_{\mathbb{F}}$'s everywhere else.

Proof: *Put $(\alpha_{uv}) = \mathbf{E}_{ij}A$. To obtain $\mathbf{E}_{ij}A$ we multiply the rows of $\mathbf{E}_{ij}$ by the columns of A. Now*

$$\alpha_{uv} = \sum_{k=1}^{n} e_{uk}a_{kv} = \sum_{k=1}^{n} \delta_{ui}\delta_{kj}a_{kv} = \delta_{ui}a_{jv}.$$

Therefore, for $u \neq i$, $\alpha_{uv} = 0_{\mathbb{F}}$, i.e., off of the i-th row the entries of $\mathbf{E}_{ij}A$ are $0_{\mathbb{F}}$, and $\alpha_{iv} = \alpha_{jv}$, that is, the i-th row of $\mathbf{E}_{ij}A$ is the j-th row of A. The case for $A\mathbf{E}_{ij}$ is similarly argued.❑

The following corollary is immediate.

108 Corollary Let $(\mathbf{E}_{ij}, \mathbf{E}_{kl}) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$, be square elementary matrices. Then

$$\mathbf{E}_{ij}\mathbf{E}_{kl} = \delta_{jk}\mathbf{E}_{il}.$$

109 Example Let $M \in \mathbf{M}_{n\times n}(\mathbb{F})$ be a matrix such that $AM = MA$ for all matrices $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Demonstrate that $M = a\mathbf{I}_n$ for some $a \in \mathbb{F}$, i.e. M is a scalar matrix.

Solution: ► *Assume* $(s,t) \in \{1,2,\ldots,n\}^2$. *Let* $M = (m_{ij})$ *and* $\mathbf{E}_{st} \in \mathbf{M}_{n\times n}(\mathbb{F})$. *Since* M *commutes with* $\mathbf{E}_{st}$ *we have*

$$\begin{bmatrix} 0 & 0 & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots \\ m_{t1} & m_{t2} & \ldots & m_{tn} \\ \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix} = \mathbf{E}_{st}M = M\mathbf{E}_{st} = \begin{bmatrix} 0 & 0 & \ldots & m_{1s} & \ldots & 0 \\ 0 & 0 & \vdots & m_{2s} & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & m_{(n-1)s} & \vdots & 0 \\ 0 & 0 & \vdots & m_{ns} & \vdots & 0 \end{bmatrix}$$

For arbitrary $s \neq t$ *we have shown that* $m_{st} = m_{ts} = 0$, *and that* $m_{ss} = m_{tt}$. *Thus the entries off the main diagonal are zero and the diagonal entries are all equal to one another, whence* M *is a scalar matrix.* ◄

110 Definition Let $\lambda \in \mathbb{F}$ and $\mathbf{E}_{ij} \in \mathbf{M}_{n\times n}(\mathbb{F})$. A square matrix in $\mathbf{M}_{n\times n}(\mathbb{F})$ of the form $\mathbf{I}_n + \lambda\mathbf{E}_{ij}$ is called a *transvection*.

111 Example The matrix

$$T = \mathbf{I}_3 + 4\mathbf{E}_{13} = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is a transvection. Observe that if

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix}$$

then

$$TA = \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 9 & 13 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix},$$

that is, pre-multiplication by T adds 4 times the third row of A to the first row of A. Similarly,

$$AT = \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 5 \\ 5 & 6 & 27 \\ 1 & 2 & 7 \end{bmatrix},$$

that is, post-multiplication by T adds 4 times the first column of A to the third row of A.

In general, we have the following theorem.

112 Theorem (Multiplication by a Transvection Matrix) Let $\mathbf{I}_n + \lambda\mathbf{E}_{ij} \in \mathbf{M}_{n\times n}(\mathbb{F})$ be a transvection and let $A \in \mathbf{M}_{n\times m}(\mathbb{F})$. Then $(\mathbf{I}_n + \lambda\mathbf{E}_{ij})A$ adds the j-th row of A to its i-th row and leaves the other rows unchanged. Similarly, if $B \in \mathbf{M}_{p\times n}(\mathbb{F})$, $B(\mathbf{I}_n + \lambda\mathbf{E}_{ij})$ adds the i-th column of B to the j-th column and leaves the other columns unchanged.

> **Proof:** *Simply observe that* $(\mathbf{I}_n + \lambda\mathbf{E}_{ij})A = A + \lambda\mathbf{E}_{ij}A$ *and* $A(\mathbf{I}_n + \lambda\mathbf{E}_{ij}) = A + \lambda A\mathbf{E}_{ij}$ *and apply Theorem 107.* ❑

Observe that the particular transvection $\mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii} \in \mathbf{M}_{n\times n}(\mathbb{F})$ consists of a diagonal matrix with $1_{\mathbb{F}}$'s everywhere on the diagonal, except on the ii-th position, where it has a λ.

113 Definition If $\lambda \neq 0_{\mathbb{F}}$, we call the matrix $\mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii}$ a *dilatation matrix*.

114 Example The matrix

$$S = \mathbf{I}_3 + (4-1)\mathbf{E}_{11} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is a dilatation matrix. Observe that if

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix}$$

then

$$SA = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 4 & 4 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix},$$

that is, pre-multiplication by S multiplies by 4 the first row of A. Similarly,

$$AS = \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 1 & 1 \\ 20 & 6 & 7 \\ 4 & 2 & 3 \end{bmatrix},$$

that is, post-multiplication by S multiplies by 4 the first column of A.

115 Theorem (Multiplication by a Dilatation Matrix) Pre-multiplication of the matrix$A \in \mathbf{M}_{n\times m}(\mathbb{F})$ by the dilatation matrix $\mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii} \in \mathbf{M}_{n\times n}(\mathbb{F})$ multiplies the i-th row of A by λ and leaves the other rows of A unchanged. Similarly, if $B \in \mathbf{M}_{p\times n}(\mathbb{F})$ post-multiplication of B by $\mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii}$ multiplies the i-th column of B by λ and leaves the other columns of B unchanged.

Proof: *This follows by direct application of Theorem 112.* ❑

116 Definition We write $\mathbf{I}_n^{ij}$ for the matrix which permutes the i-th row with the j-th row of the identity matrix. We call $\mathbf{I}_n^{ij}$ a *transposition matrix.*

117 Example We have

$$\mathbf{I}_4^{(23)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

If

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix},$$

then

$$\mathbf{I}_4^{(23)}A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \\ 13 & 14 & 15 & 16 \end{bmatrix},$$

and

$$A\mathbf{I}_4^{(23)} = \begin{bmatrix} 1 & 3 & 2 & 4 \\ 5 & 7 & 6 & 8 \\ 9 & 11 & 10 & 12 \\ 13 & 15 & 14 & 16 \end{bmatrix}.$$

118 Theorem (Multiplication by a Transposition Matrix) If $A \in \mathbf{M}_{n\times m}(\mathbb{F})$, then $\mathbf{I}_n^{ij}A$ is the matrix obtained from A permuting the the i-th row with the j-th row of A. Similarly, if $B \in \mathbf{M}_{p\times n}(\mathbb{F})$, then $B\mathbf{I}_n^{ij}$ is the matrix obtained from B by permuting the i-th column with the j-th column of B.

Proof: *We must prove that $\mathbf{I}_n^{ij}A$ exchanges the i-th and j-th rows but leaves the other rows unchanged. But this follows upon observing that*

$$\mathbf{I}_n^{ij} = \mathbf{I}_n + \mathbf{E}_{ij} + \mathbf{E}_{ji} - \mathbf{E}_{ii} - \mathbf{E}_{jj}$$

and appealing to Theorem 107.

❑

119 Definition A square matrix which is either a transvection matrix, a dilatation matrix or a transposition matrix is called an *elimination matrix.*

☞ *In a very loose way, we may associate pre-multiplication of a matrix* A *by another matrix with an operation on the rows of* A*, and post-multiplication of a matrix* A *by another with an operation on the columns of* A*.*

Homework

Problem 2.4.1 Consider the matrices

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 4 & -2 & 4 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix}.$$

Find a specific dilatation matrix D, a specific transposition matrix P, and a specific transvection matrix T such that $B = TDAP$.

Problem 2.4.2 The matrix

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is transformed into the matrix

$$B = \begin{bmatrix} h-g & g & i \\ e-d & d & f \\ 2b-2a & 2a & 2c \end{bmatrix}$$

by a series of row and column operations. Find explicit permutation matrices P, P', an explicit dilatation matrix D, and an explicit transvection matrix T such that

$$B = DPAP'T.$$

Problem 2.4.3 Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Prove that if

$$(\forall X \in \mathbf{M}_{n\times n}(\mathbb{F})), \ (\mathbf{tr}(AX) = \mathbf{tr}(BX)),$$

then $A = B$.

Problem 2.4.4 Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$ be such that

$$(\forall X \in \mathbf{M}_{n\times n}(\mathbb{R})), \ ((XA)^2 = \mathbf{0}_n).$$

Prove that $A = \mathbf{0}_n$.

2.5 Matrix Inversion

120 Definition Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$. Then A is said to be *left-invertible* if $\exists L \in \mathbf{M}_{n\times m}(\mathbb{F})$ such that $LA = \mathbf{I}_n$. A is said to be *right-invertible* if $\exists R \in \mathbf{M}_{n\times m}(\mathbb{F})$ such that $AR = \mathbf{I}_m$. A matrix is said to be *invertible* if it possesses a right and a left inverse. A matrix which is not invertible is said to be *singular.*

121 Example The matrix $A \in \mathbf{M}_{2\times 3}(\mathbb{R})$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

has infinitely many right-inverses of the form

$$R_{(x,y)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ x & y \end{bmatrix}.$$

For

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ x & y \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

regardless of the values of x and y. Observe, however, that A does not have a left inverse, for

$$\begin{bmatrix} a & b \\ c & d \\ f & g \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ f & g & 0 \end{bmatrix},$$

which will never give $\mathbf{I}_3$ regardless of the values of a, b, c, d, f, g.

122 Example If $\lambda \neq 0$, then the scalar matrix $\lambda \mathbf{I}_n$ is invertible, for

$$(\lambda \mathbf{I}_n)\left(\lambda^{-1}\mathbf{I}_n\right) = \mathbf{I}_n = \left(\lambda^{-1}\mathbf{I}_n\right)(\lambda \mathbf{I}_n).$$

123 Example The zero matrix $\mathbf{0}_n$ is singular.

124 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ a square matrix possessing a left inverse L and a right inverse R. Then $L = R$. Thus an invertible square matrix possesses a unique inverse.

Proof: *Observe that we have* $LA = \mathbf{I}_n = AR$. *Then*

$$L = L\mathbf{I}_n = L(AR) = (LA)R = \mathbf{I}_n R = R.$$

❑

125 Definition The subset of $\mathbf{M}_{n\times n}(\mathbb{F})$ of all invertible $n \times n$ matrices is denoted by $\mathbf{GL}_n(\mathbb{F})$, read "the linear group of rank n over $\mathbb{F}$."

126 Corollary Let $(A, B) \in (\mathbf{GL}_n(\mathbb{F}))^2$. Then AB is also invertible and

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

Proof: *Since* AB *is a square matrix, it suffices to notice that*

$$B^{-1}A^{-1}(AB) = (AB)B^{-1}A^{-1} = \mathbf{I}_n$$

and that since the inverse of a square matrix is unique, we must have $B^{-1}A^{-1} = (AB)^{-1}$. ❑

127 Corollary If a square matrix $S \in \mathbf{M}_{n\times n}(\mathbb{F})$ is invertible, then S^{-1} is also invertible and $(S^{-1})^{-1} = S$, in view of the uniqueness of the inverses of square matrices.

128 Corollary If a square matrix $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ is invertible, then A^T is also invertible and $(A^T)^{-1} = (A^{-1})^T$.

Proof: *We claim that* $(A^T)^{-1} = (A^{-1})^T$. *For*

$$AA^{-1} = \mathbf{I}_n \implies (AA^{-1})^T = \mathbf{I}_n^T \implies (A^{-1})^T A^T = \mathbf{I}_n,$$

where we have used Theorem 91. ❑

The next few theorems will prove that elimination matrices are invertible matrices.

129 Theorem (Invertibility of Transvections) Let $\mathbf{I}_n + \lambda\mathbf{E}_{ij} \in \mathbf{M}_{n\times n}(\mathbb{F})$ be a transvection, and let $i \neq j$. Then

$$(\mathbf{I}_n + \lambda\mathbf{E}_{ij})^{-1} = \mathbf{I}_n - \lambda\mathbf{E}_{ij}.$$

Proof: *Expanding the product*

$$\begin{aligned}(\mathbf{I}_n + \lambda\mathbf{E}_{ij})(\mathbf{I}_n - \lambda\mathbf{E}_{ij}) &= \mathbf{I}_n + \lambda\mathbf{E}_{ij} - \lambda\mathbf{E}_{ij} - \lambda^2\mathbf{E}_{ij}\mathbf{E}_{ij} \\ &= \mathbf{I}_n - \lambda^2\delta_{ij}\mathbf{E}_{ij} \\ &= \mathbf{I}_n,\end{aligned}$$

since $i \neq j$. ❑

130 Example By Theorem 129, we have

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

131 Theorem (Invertibility of Dilatations) Let $\lambda \neq 0_{\mathbb{F}}$. Then

$$(\mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii})^{-1} = \mathbf{I}_n + (\lambda^{-1} - 1_{\mathbb{F}})\mathbf{E}_{ii}.$$

Proof: *Expanding the product*

$$\begin{aligned}(\mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii})(\mathbf{I}_n + (\lambda^{-1} - 1_{\mathbb{F}})\mathbf{E}_{ii}) &= \mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii} \\ &\quad +(\lambda^{-1} - 1_{\mathbb{F}})\mathbf{E}_{ii} \\ &\quad\quad +(\lambda - 1_{\mathbb{F}})(\lambda^{-1} - 1_{\mathbb{F}})\mathbf{E}_{ii} \\ &= \mathbf{I}_n + (\lambda - 1_{\mathbb{F}})\mathbf{E}_{ii} \\ &\quad +(\lambda^{-1} - 1_{\mathbb{F}})\mathbf{E}_{ii} \\ &\quad\quad +(\lambda - 1_{\mathbb{F}})(\lambda^{-1} - 1_{\mathbb{F}}))\mathbf{E}_{ii} \\ &= \mathbf{I}_n + (\lambda - 1_{\mathbb{F}} + \lambda^{-1} - 1_{\mathbb{F}} + 1_{\mathbb{F}} \\ &\quad\quad -\lambda - \lambda^{-1} - 1_{\mathbb{F}}))\mathbf{E}_{ii} \\ &= \mathbf{I}_n,\end{aligned}$$

proving the assertion. ❑

132 Example By Theorem 131, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Repeated applications of Theorem 131 gives the following corollary.

133 Corollary If $\lambda_1\lambda_2\lambda_3 \cdots \lambda_n \neq 0_{\mathbb{F}}$, then

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

is invertible and

$$\begin{bmatrix} \lambda_1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}^{-1} = \begin{bmatrix} \lambda_1^{-1} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^{-1} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_3^{-1} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda_n^{-1} \end{bmatrix}$$

134 Theorem (Invertibility of Permutation Matrices) Let $\tau \in S_n$ be a permutation. Then

$$(\mathbf{I}_n^{ij})^{-1} = (\mathbf{I}_n^{ij})^T.$$

Proof: *By Theorem 118 pre-multiplication of $\mathbf{I}_n^{ij}$ by $\mathbf{I}_n^{ij}$ exchanges the i-th row with the j-th row, meaning that they return to the original position in $\mathbf{I}_n$. Observe in particular that $\mathbf{I}_n^{ij} = (\mathbf{I}_n^{ij})^T$, and so $\mathbf{I}_n^{ij}(\mathbf{I}_n^{ij})^T = \mathbf{I}_n$.* ❑

135 Example By Theorem 134, we have

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

136 Corollary If a square matrix can be represented as the product of elimination matrices of the same size, then it is invertible.

Proof: *This follows from Corollary 126, and Theorems 129, 131, and 134.* ❑

137 Example Observe that

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

is the transvection $\mathbf{I}_3 + 4\mathbf{E}_{23}$ followed by the dilatation of the second column of this transvection by 3. Thus

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and so

$$\begin{aligned} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}^{-1} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & -\frac{4}{3} \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

138 Example We have

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

hence

$$\begin{bmatrix}1 & 1 & 1\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}^{-1} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{bmatrix}^{-1}\begin{bmatrix}1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}^{-1}$$

$$= \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & -1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$= \begin{bmatrix}1 & -1 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{bmatrix}.$$

In the next section we will give a general method that will permit us to find the inverse of a square matrix when it exists.

139 Example Let $T = \begin{bmatrix}a & b\\ c & d\end{bmatrix} \in \mathbf{M}_{2\times 2}(\mathbb{R})$. Then

$$\begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix} = (ad - bc)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$$

Thus if $ad - bc \neq 0$ we see that

$$T^{-1} = \begin{bmatrix}\frac{d}{ad-bc} & -\frac{b}{ad-bc}\\ -\frac{c}{ad-bc} & \frac{a}{ad-bc}\end{bmatrix}.$$

140 Example If

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

then $\mathbf{A}$ is invertible, for an easy computation shews that

$$\mathbf{A}^2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}^2 = 4\mathbf{I}_4,$$

whence the inverse sought is

$$\mathbf{A}^{-1} = \frac{1}{4}\mathbf{A} = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/4 & -1/4 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ 1/4 & -1/4 & -1/4 & 1/4 \end{bmatrix}.$$

141 Example A matrix $\mathbf{A} \in \mathbf{M}_{n\times n}(\mathbb{R})$ is said to be *nilpotent* of index k if satisfies $\mathbf{A} \neq \mathbf{0}_n, \mathbf{A}^2 \neq \mathbf{0}_n, \ldots, \mathbf{A}^{k-1} \neq \mathbf{0}_n$ and $\mathbf{A}^k = \mathbf{0}_n$ for integer $k \geq 1$. Prove that if $\mathbf{A}$ is nilpotent, then $\mathbf{I}_n - \mathbf{A}$ is invertible and find its inverse.

Solution: ► *To motivate the solution, think that instead of a matrix, we had a real number x with $|x| < 1$. Then the inverse of $1 - x$ is*

$$(1-x)^{-1} = \frac{1}{1-x} = 1 + x + x^2 + x^3 + \cdots.$$

Notice now that since $\mathbf{A}^k = \mathbf{0}_n$, then $\mathbf{A}^p = \mathbf{0}_n$ for $p \geq k$. We conjecture thus that

$$(\mathbf{I}_n - \mathbf{A})^{-1} = \mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1}.$$

The conjecture is easily verified, as

$$\begin{aligned}(\mathbf{I}_n - \mathbf{A})(\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1}) &= \mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1} \\ &\quad -(\mathbf{A} + \mathbf{A}^2 + \mathbf{A}^3 + \cdots + \mathbf{A}^k) \\ &= \mathbf{I}_n\end{aligned}$$

and

$$\begin{aligned}(\mathbf{I}_n + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1})(\mathbf{I}_n - \mathbf{A}) &= \mathbf{I}_n - \mathbf{A} + \mathbf{A} - \mathbf{A}^2 + \mathbf{A}^3 - \mathbf{A}^4 + \cdots \\ &\quad \cdots + \mathbf{A}^{k-2} - \mathbf{A}^{k-1} + \mathbf{A}^{k-1} - \mathbf{A}^k \\ &= \mathbf{I}_n.\end{aligned}$$

◄

142 Example The inverse of $A \in \mathbf{M}_{3\times 3}(\mathbb{Z}_5)$,

$$A = \begin{bmatrix} \overline{2} & \overline{0} & \overline{0} \\ \overline{0} & \overline{3} & \overline{0} \\ \overline{0} & \overline{0} & \overline{4} \end{bmatrix}$$

is

$$A^{-1} = \begin{bmatrix} \overline{3} & \overline{0} & \overline{0} \\ \overline{0} & \overline{2} & \overline{0} \\ \overline{0} & \overline{0} & \overline{4} \end{bmatrix},$$

as

$$AA^{-1} = \begin{bmatrix} \overline{2} & \overline{0} & \overline{0} \\ \overline{0} & \overline{3} & \overline{0} \\ \overline{0} & \overline{0} & \overline{4} \end{bmatrix} \begin{bmatrix} \overline{3} & \overline{0} & \overline{0} \\ \overline{0} & \overline{2} & \overline{0} \\ \overline{0} & \overline{0} & \overline{4} \end{bmatrix} = \begin{bmatrix} \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{1} & \overline{0} \\ \overline{0} & \overline{0} & \overline{1} \end{bmatrix}$$

143 Example (Putnam Exam, 1991) Let A and B be different $n \times n$ matrices with real entries. If $A^3 = B^3$ and $A^2B = B^2A$, prove that $A^2 + B^2$ is not invertible.

Solution: ► *Observe that*

$$(A^2 + B^2)(A - B) = A^3 - A^2B + B^2A - B^3 = \mathbf{0}_n.$$

If $A^2 + B^2$ were invertible, then we would have

$$A - B = (A^2 + B^2)^{-1}(A^2 + B^2)(A - B) = \mathbf{0}_n,$$

contradicting the fact that A and B are different matrices. ◄

144 Lemma If $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ has a row or a column consisting all of $0_{\mathbb{F}}$'s, then A is singular.

Proof: *If A were invertible, the (i,i)-th entry of the product of its inverse with A would be $1_{\mathbb{F}}$. But if the i-th row of A is all $0_{\mathbb{F}}$'s, then $\sum_{k=1}^{n} a_{ik}b_{ki} = 0_{\mathbb{F}}$, so the (i,i) entry of any matrix product with A is $0_{\mathbb{F}}$, and never $1_{\mathbb{F}}$.* ❑

Problem 2.5.1 The inverse of the matrix $A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}$ is the matrix $A^{-1} = \begin{bmatrix} a & 1 & -1 \\ 1 & b & 1 \\ -1 & 1 & 0 \end{bmatrix}$. Determine a and b.

Problem 2.5.2 A square matrix A satisfies $A^3 \neq \mathbf{0}_n$ but $A^4 = \mathbf{0}_n$. Demonstrate that $\mathbf{I}_n + A$ is invertible and find, with proof, its inverse.

Problem 2.5.3 Prove or disprove! If $(A, B, A + B) \in (\mathbf{GL}_n(\mathbb{R}))^3$ then $(A + B)^{-1} = A^{-1} + B^{-1}$.

Problem 2.5.4 Let $S \in \mathbf{GL}_n(\mathbb{F})$, $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$, and k a positive integer. Prove that if $B = SAS^{-1}$ then $B^k = SA^kS^{-1}$.

Problem 2.5.5 Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ and let k be a positive integer. Prove that A is invertible if and only if A^k is invertible.

Problem 2.5.6 Let $S \in \mathbf{GL}_n(\mathbb{C})$, $A \in \mathbf{M}_{n\times n}(\mathbb{C})$ with $A^k = \mathbf{0}_n$ for some positive integer k. Prove that both $\mathbf{I}_n - SAS^{-1}$ and $\mathbf{I}_n - S^{-1}AS$ are invertible and find their inverses.

Problem 2.5.7 Let A and B be square matrices of the same size such that both $A - B$ and $A + B$ are invertible. Put $C = (A - B)^{-1} + (A + B)^{-1}$. Prove that

$$ACA - ACB + BCA - BCB = 2A.$$

Problem 2.5.8 Let A, B, C be non-zero square matrices of the same size over the same field and such that $ABC = \mathbf{0}_n$. Prove that at least two of these three matrices are not invertible.

Problem 2.5.9 Let $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$ be such that $A^2 = B^2 = (AB)^2 = \mathbf{I}_n$. Prove that $AB = BA$.

Problem 2.5.10 Let $A = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ b & b & a & \cdots & b \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ b & b & b & \cdots & a \end{bmatrix} \in \mathbf{M}_{n\times n}(\mathbb{F})$, $n > 1$, $(a, b) \in \mathbb{F}^2$. Determine when A is invertible and find this inverse when it exists.

Problem 2.5.11 Let $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$ be matrices such that $A + B = AB$. Demonstrate that $A - \mathbf{I}_n$ is invertible and find this inverse.

Problem 2.5.12 Let $S \in \mathbf{GL}_n(\mathbb{F})$ and $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Prove that $\mathbf{tr}\,(A) = \mathbf{tr}\,(SAS^{-1})$.

Problem 2.5.13 Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$ be a skew-symmetric matrix. Prove that $\mathbf{I}_n + A$ is invertible. Furthermore, if $B = (\mathbf{I}_n - A)(\mathbf{I}_n + A)^{-1}$, prove that $B^{-1} = B^T$.

Problem 2.5.14 A matrix $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ is said to be a *magic square* if the sum of each individual row equals the sum of each individual column. Assume that A is a magic square and invertible. Prove that A^{-1} is also a magic square.

2.6 Block Matrices

145 Definition Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, $B \in \mathbf{M}_{m\times s}(\mathbb{F})$, $C \in \mathbf{M}_{r\times n}(\mathbb{F})$, $D \in \mathbf{M}_{r\times s}(\mathbb{F})$. We use the notation

$$L = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

for the *block matrix* $L \in \mathbf{M}_{(m+r)\times(n+s)}(\mathbb{F})$.

If $(A, A') \in (\mathbf{M}_m(\mathbb{F}))^2$, $(B, B') \in (\mathbf{M}_{m\times n}(\mathbb{F}))^2$, $(C, C') \in (\mathbf{M}_{n\times m}(\mathbb{F}))^2$, $(D, D') \in (\mathbf{M}_m(\mathbb{F}))^2$, *and*

$$S = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad T = \left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right],$$

then it is easy to verify that

$$ST = \left[\begin{array}{c|c} AA' + BC' & AB' + BD' \\ \hline CA' + DC' & CB' + DD' \end{array}\right].$$

146 Lemma Let $L \in \mathbf{M}_{(m+r)\times(m+r)}(\mathbb{F})$ be the square block matrix

$$L = \left[\begin{array}{c|c} A & C \\ \hline \mathbf{0}_{r\times m} & B \end{array}\right],$$

with square matrices $A \in \mathbf{M}_m(\mathbb{F})$ and $B \in \mathbf{M}_{r\times r}(\mathbb{F})$, and a matrix $C \in \mathbf{M}_{m\times r}(\mathbb{F})$. Then L is invertible if and only if A and B are, in which case

$$L^{-1} = \left[\begin{array}{c|c} A^{-1} & -A^{-1}CB^{-1} \\ \hline \mathbf{0}_{r\times m} & B^{-1} \end{array}\right]$$

Proof: *Assume first that* A, *and* B *are invertible. Direct calculation yields*

$$\begin{aligned}
\left[\begin{array}{c|c} A & C \\ \hline \mathbf{0}_{r\times m} & B \end{array}\right]\left[\begin{array}{c|c} A^{-1} & -A^{-1}CB^{-1} \\ \hline \mathbf{0}_{r\times m} & B^{-1} \end{array}\right] &= \left[\begin{array}{c|c} AA^{-1} & -AA^{-1}CB^{-1} + CB^{-1} \\ \hline \mathbf{0}_{r\times m} & BB^{-1} \end{array}\right] \\
&= \left[\begin{array}{c|c} \mathbf{I}_m & \mathbf{0}_{m\times r} \\ \hline \mathbf{0}_{r\times m} & \mathbf{I}_r \end{array}\right] \\
&= \mathbf{I}_{m+r}.
\end{aligned}$$

Assume now that L *is invertible,* $L^{-1} = \left[\begin{array}{c|c} E & H \\ \hline J & K \end{array}\right]$, *with* $E \in \mathbf{M}_m(\mathbb{F})$ *and* $K \in \mathbf{M}_{r\times r}(\mathbb{F})$, *but that, say,* B *is singular. Then*

$$\begin{aligned}
\left[\begin{array}{c|c} \mathbf{I}_m & \mathbf{0}_{m\times r} \\ \hline \mathbf{0}_{r\times m} & \mathbf{I}_r \end{array}\right] &= LL^{-1} \\
&= \left[\begin{array}{c|c} A & C \\ \hline \mathbf{0}_{r\times m} & B \end{array}\right]\left[\begin{array}{c|c} E & H \\ \hline J & K \end{array}\right] \\
&= \left[\begin{array}{c|c} AE + CJ & AH + BK \\ \hline BJ & BK \end{array}\right],
\end{aligned}$$

which gives $BK = \mathbf{I}_r$, *i.e.,* B *is invertible, a contradiction.* ❑

2.7 Rank of a Matrix

147 Definition Let $(A, B) \in (\mathbf{M}_{m\times n}(\mathbb{F}))^2$. We say that A is *row-equivalent* to B if there exists a matrix $R \in \mathbf{GL}_m(\mathbb{F})$ such that $B = RA$. Similarly, we say that A is *column-equivalent* to B if there exists a matrix $C \in \mathbf{GL}_m(\mathbb{F})$ such that $B = AC$. We say that A and B are *equivalent* if $\exists(P, Q) \in \mathbf{GL}_m(\mathbb{F}) \times \mathbf{GL}_n(\mathbb{F})$ such that $B = PAQ$.

148 Theorem Row equivalence, column equivalence, and equivalence are equivalence relations.

Proof: *We prove the result for row equivalence. The result for column equivalence, and equivalence are analogously proved.*

Since $\mathbf{I}_m \in \mathbf{GL}_m(\mathbb{F})$ *and* $A = \mathbf{I}_m A$*, row equivalence is a reflexive relation. Assume* $(A, B) \in (\mathbf{M}_{m\times n}(\mathbb{F}))^2$ *and that* $\exists P \in \mathbf{GL}_m(\mathbb{F})$ *such that* $B = PA$*. Then* $A = P^{-1}B$ *and since* $P^{-1} \in \mathbf{GL}_m(\mathbb{F})$*, we see that row equivalence is a symmetric relation. Finally assume* $(A, B, C) \in (\mathbf{M}_{m\times n}(\mathbb{F}))^3$ *and that* $\exists P \in \mathbf{GL}_m(\mathbb{F})$*,* $\exists P' \in \mathbf{GL}_m(\mathbb{F})$ *such that* $A = PB$*,* $B = P'C$*. Then* $A = PP'C$*. But* $PP' \in \mathbf{GL}_m(\mathbb{F})$ *in view of Corollary 126. This completes the proof.* ❑

149 Theorem Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$. Then A can be reduced, by means of pre-multiplication and post-multiplication by elimination matrices, to a unique matrix of the form

$$D_{m,n,r} = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(n-r)} \\ \hline \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right], \tag{2.16}$$

called the *Hermite normal form* of A. Thus there exist $P \in \mathbf{GL}_m(\mathbb{F})$, $Q \in \mathbf{GL}_n(\mathbb{F})$ such that $D_{m,n,r} = PAQ$. The integer $r \geq 0$ is called the *rank* of the matrix A which we denote by **rank** (A).

Proof: *If* A *is the* $m \times n$ *zero matrix, then the theorem is obvious, taking* $r = 0$*. Assume hence that* A *is not the zero matrix. We proceed as follows using the* Gauß-Jordan Algorithm.

GJ-1 *Since* A *is a non-zero matrix, it has a non-zero column. By means of permutation matrices we move this column to the first column.*

GJ-2 *Since this column is a non-zero column, it must have an entry* $a \neq 0_{\mathbb{F}}$*. Again, by means of permutation matrices, we move the row on which this entry is to the first row.*

GJ-3 *By means of a dilatation matrix with scale factor* a^{-1}*, we make this new* $(1, 1)$ *entry into a* $1_{\mathbb{F}}$*.*

GJ-4 *By means of transvections (adding various multiples of row* 1 *to the other rows) we now annihilate every entry below the entry* $(1, 1)$*.*

This process ends up in a matrix of the form

$$P_1AQ_1 = \left[\begin{array}{c|cccc} 1_{\mathbb{F}} & * & * & \cdots & * \\ \hline 0_{\mathbb{F}} & b_{22} & b_{23} & \cdots & b_{2n} \\ 0_{\mathbb{F}} & b_{32} & b_{33} & \cdots & b_{3n} \\ 0_{\mathbb{F}} & \vdots & \vdots & \cdots & \vdots \\ 0_{\mathbb{F}} & b_{m2} & b_{m3} & \cdots & b_{mn} \end{array}\right]. \tag{2.17}$$

Here the asterisks represent unknown entries. Observe that the b*'s form a* $(m-1)\times(n-1)$ *matrix.*

GJ-5 *Apply* **GJ-1** *through* **GJ-4** *to the matrix of the* b*'s.*

Observe that this results in a matrix of the form

$$P_2AQ_2 = \left[\begin{array}{cc|ccc} 1_{\mathbb{F}} & * & * & \cdots & * \\ 0_{\mathbb{F}} & 1_{\mathbb{F}} & * & \cdots & * \\ \hline 0_{\mathbb{F}} & 0_{\mathbb{F}} & c_{33} & \cdots & c_{3n} \\ 0_{\mathbb{F}} & \vdots & \vdots & \cdots & \vdots \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & c_{m3} & \cdots & c_{mn} \end{array}\right]. \tag{2.18}$$

GJ-6 *Add the appropriate multiple of column 1 to column 2, that is, apply a transvection, in order to make the entry in the* $(1, 2)$ *position* $0_{\mathbb{F}}$.

This now gives a matrix of the form

$$P_3AQ_3 = \left[\begin{array}{cc|ccc} 1_{\mathbb{F}} & 0_{\mathbb{F}} & * & \cdots & * \\ 0_{\mathbb{F}} & 1_{\mathbb{F}} & * & \cdots & * \\ \hline 0_{\mathbb{F}} & 0_{\mathbb{F}} & c_{33} & \cdots & c_{3n} \\ 0_{\mathbb{F}} & \vdots & \vdots & \cdots & \vdots \\ 0_{\mathbb{F}} & 0_{\mathbb{F}} & c_{m3} & \cdots & c_{mn} \end{array}\right]. \tag{2.19}$$

The matrix of the c*'s has size* $(m-2) \times (n-2)$.

GJ-7 *Apply* **GJ-1** *through* **GJ-6** *to the matrix of the* c*'s, etc.*

Observe that this process eventually stops, and in fact, it is clear that $\mathbf{rank}\,(A) \leq \mathbf{min}(m, n)$.

Suppose now that A *were equivalent to a matrix* $D_{m,n,s}$ *with* $s > r$. *Since matrix equivalence is an equivalence relation,* $D_{m,n,s}$ *and* $D_{m,n,r}$ *would be equivalent, and so there would be* $R \in \mathbf{GL}_m(\mathbb{F})$, $S \in \mathbf{GL}_n(\mathbb{F})$, *such that* $RD_{m,n,r}S = D_{m,n,s}$, *that is,* $RD_{m,n,r} = D_{m,n,s}S^{-1}$. *Partition* R *and* S^{-1} *as follows*

$$R = \left[\begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{21} & R_{22} \end{array}\right], \quad S^{-1} = \left[\begin{array}{c|cc} S_{11} & S_{12} & S_{13} \\ \hline S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{array}\right],$$

with $(R_{11}, S_{11})^2 \in (\mathbf{M}_{r\times r}(\mathbb{F}))^2$, $S_{22} \in \mathbf{M}_{(s-r)\times(s-r)}(\mathbb{F})$. *We have*

$$RD_{m,n,r} = \left[\begin{array}{c|c} R_{11} & R_{12} \\ \hline R_{21} & R_{22} \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{(m-r)\times r} \\ \hline \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{r\times(m-r)} \end{array}\right] = \left[\begin{array}{c|c} R_{11} & \mathbf{0}_{(m-r)\times r} \\ \hline R_{21} & \mathbf{0}_{r\times(m-r)} \end{array}\right],$$

and

$$D_{m,n,s}S^{-1} = \left[\begin{array}{c|cc} \mathbf{I}_r & \mathbf{0}_{r\times(s-r)} & \mathbf{0}_{r\times(n-s)} \\ \hline \mathbf{0}_{(s-r)\times r} & \mathbf{I}_{s-r} & \mathbf{0}_{(s-r)\times(n-s)} \\ \mathbf{0}_{(m-s)\times r} & \mathbf{0}_{(m-s)\times(s-r)} & \mathbf{0}_{(m-s)\times(n-s)} \end{array}\right] \left[\begin{array}{c|cc} S_{11} & S_{12} & S_{13} \\ \hline S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{array}\right]$$

$$= \left[\begin{array}{c|cc} S_{11} & S_{12} & S_{13} \\ \hline S_{21} & S_{22} & S_{23} \\ \mathbf{0}_{(m-s)\times r} & \mathbf{0}_{(m-s)\times(s-r)} & \mathbf{0}_{(m-s)\times(n-s)} \end{array}\right].$$

Since we are assuming

$$\left[\begin{array}{c|c} R_{11} & \mathbf{0}_{(m-r)\times r} \\ \hline R_{21} & \mathbf{0}_{r\times(m-r)} \end{array}\right] = \left[\begin{array}{c|cc} S_{11} & S_{12} & S_{13} \\ \hline S_{21} & S_{22} & S_{23} \\ \mathbf{0}_{(m-s)\times r} & \mathbf{0}_{(m-s)\times(s-r)} & \mathbf{0}_{(m-s)\times(n-s)} \end{array}\right],$$

we must have $S_{12} = \mathbf{0}_{r\times(s-r)}$, $S_{13} = \mathbf{0}_{r\times(n-s)}$, $S_{22} = \mathbf{0}_{(s-r)\times(s-r)}$, $S_{23} = \mathbf{0}_{(s-r)\times(n-s)}$. *Hence*

$$S^{-1} = \left[\begin{array}{c|cc} S_{11} & \mathbf{0}_{r\times(s-r)} & \mathbf{0}_{r\times(n-s)} \\ \hline S_{21} & \mathbf{0}_{(s-r)\times(s-r)} & \mathbf{0}_{(s-r)\times(n-s)} \\ S_{31} & S_{32} & S_{33} \end{array}\right].$$

The matrix

$$\left[\begin{array}{c|c} \mathbf{0}_{(s-r)\times(s-r)} & \mathbf{0}_{(s-r)\times(n-s)} \\ \hline S_{32} & S_{33} \end{array}\right]$$

is non-invertible, by virtue of Lemma 144. This entails that S^{-1} *is non-invertible by virtue of Lemma 146. This is a contradiction, since* S *is assumed invertible, and hence* S^{-1} *must also be invertible.* ❑

☞ *Albeit the rank of a matrix is unique, the matrices* $\mathbf{P}$ *and* $\mathbf{Q}$ *appearing in Theorem 149 are not necessarily unique. For example, the matrix*

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

has rank 2*, the matrix*

$$\begin{bmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}$$

is invertible, and an easy computation shews that

$$\begin{bmatrix} 1 & 0 & x \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix},$$

regardless of the values of x *and* y.

150 Corollary Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$. Then $\mathbf{rank}\,(A) = \mathbf{rank}\,(A^T)$.

Proof: *Let* $P, Q, D_{m,n,r}$ *as in Theorem 149. Observe that* P^T, Q^T *are invertible. Then*

$$PAQ = D_{m,n,r} \implies Q^TA^TP^T = D_{m,n,r}^T = D_{n,m,r},$$

and since this last matrix has the same number of $1_{\mathbb{F}}$*'s as* $D_{m,n,r}$*, the corollary is proven.* ❑

151 Example Shew that

$$A = \begin{bmatrix} 0 & 2 & 3 \\ 0 & 1 & 0 \end{bmatrix}$$

has $\mathbf{rank}\,(A) = 2$ and find invertible matrices $P \in \mathbf{GL}_2(\mathbb{R})$ and $Q \in \mathbf{GL}_3(\mathbb{R})$ such that

$$PAQ = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

Solution: ▶ *We first transpose the first and third columns by effecting*

$$\begin{bmatrix} 0 & 2 & 3 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

We now subtract twice the second row from the first, by effecting

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

Finally, we divide the first row by 3,

$$\begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

We conclude that

$$\begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 2 & 3 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

from where we may take

$$P = \begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 \\ 0 & 1 \end{bmatrix}$$

and

$$Q = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

◀

In practice it is easier to do away with the multiplication by elimination matrices and perform row and column operations on the *augmented* $(m + n) \times (m + n)$ *matrix*

$$\left[\begin{array}{c|c} \mathbf{I}_n & \mathbf{0}_{n\times m} \\ \hline A & \mathbf{I}_m \end{array}\right].$$

152 Definition Denote the rows of a matrix $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ by $R_1, R_2, \ldots, R_m$, and its columns by $C_1, C_2, \ldots, C_n$. The elimination operations will be denoted as follows.

- Exchanging the i-th row with the j-th row, which we denote by $R_i \leftrightarrow R_j$, and the s-th column by the t-th column by $C_s \leftrightarrow C_t$.
- A dilatation of the i-th row by a non-zero scalar $\alpha \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$, we will denote by $\alpha R_i \to R_i$. Similarly, $\beta C_j \to C_j$ denotes the dilatation of the j-th column by the non-zero scalar β.
- A transvection on the rows will be denoted by $R_i + \alpha R_j \to R_i$, and one on the columns by $C_s + \beta C_t \to C_s$.

153 Example Find the Hermite normal form of

$$A = \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

Solution: ► *First observe that* $\mathbf{rank}\,(A) \leq \mathbf{min}(4, 2) = 2$*, so the rank can be either* 1 *or* 2 *(why* *not 0?). Form the augmented matrix*

$$\left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right].$$

Perform $R_5 + R_3 \to R_5$ *and* $R_6 + R_3 \to R_6$ *successively, obtaining*

$$\left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & 0 & 0 & 1 \end{array}\right].$$

Perform $R_6 - 2R_5 \to R_6$

$$\left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & -2 & 1 \end{array}\right].$$

Perform $R_4 \leftrightarrow R_5$

$$\left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline -1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -2 & 1 \end{array}\right].$$

Finally, perform $-R_3 \to R_3$

$$\left[\begin{array}{cc|cccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -2 & 1 \end{array}\right].$$

We conclude that

$$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

◀

154 Theorem Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, $B \in \mathbf{M}_{n\times p}(\mathbb{F})$. Then

$$\mathbf{rank}\,(AB) \le \mathbf{min}(\mathbf{rank}\,(A)\,, \mathbf{rank}\,(B)).$$

Proof: *We prove that* $\mathbf{rank}\,(A) \ge \mathbf{rank}\,(AB)$. *The proof that* $\mathbf{rank}\,(B) \ge \mathbf{rank}\,(AB)$ *is similar and left to the reader. Put* $r = \mathbf{rank}\,(A)\,, s = \mathbf{rank}\,(AB)$. *There exist matrices* $P \in \mathbf{GL}_m(\mathbb{F})$, $Q \in \mathbf{GL}_n(\mathbb{F})$, $S \in \mathbf{GL}_m(\mathbb{F})$, $T \in \mathbf{GL}_p(\mathbb{F})$ *such that*

$$PAQ = D_{m,n,r}, \quad SABT = D_{m,p,s}.$$

Now

$$D_{m,p,s} = SABT = SP^{-1}D_{m,n,r}Q^{-1}BT,$$

from where it follows that

$$PS^{-1}D_{m,p,s} = D_{m,n,r}Q^{-1}BT.$$

Now the proof is analogous to the uniqueness proof of Theorem 149. Put $U = PS^{-1} \in \mathbf{GL}_m(\mathbb{R})$ *and* $V = Q^{-1}BT \in \mathbf{M}_{n\times p}(\mathbb{F})$, *and partition* U *and* V *as follows:*

$$U = \left[\begin{array}{c|c} U_{11} & U_{12} \\ \hline U_{21} & U_{22} \end{array}\right], \quad V = \left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right],$$

with $U_{11} \in \mathbf{M}_s(\mathbb{F}), V_{11} \in \mathbf{M}_{r\times r}(\mathbb{F})$. *Then*

$$UD_{m,p,s} = \left[\begin{array}{c|c} U_{11} & U_{12} \\ \hline U_{21} & U_{22} \end{array}\right] \left[\begin{array}{c|c} \mathbf{I}_s & \mathbf{0}_{s\times(p-s)} \\ \hline \mathbf{0}_{(m-s)\times s} & \mathbf{0}_{(m-s)\times(p-s)} \end{array}\right] \in \mathbf{M}_{m\times p}(\mathbb{F}),$$

and

$$D_{m,p,s}V = \left[\begin{array}{c|c} \mathbf{I}_r & \mathbf{0}_{r\times(n-r)} \\ \hline \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right]\left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline V_{21} & V_{22} \end{array}\right] \in \mathbf{M}_{m\times p}(\mathbb{F}).$$

From the equality of these two $m \times p$ matrices, it follows that

$$\left[\begin{array}{c|c} U_{11} & \mathbf{0}_{s\times(p-s)} \\ \hline U_{21} & \mathbf{0}_{(m-s)\times(p-s)} \end{array}\right] = \left[\begin{array}{c|c} V_{11} & V_{12} \\ \hline \mathbf{0}_{(m-r)\times r} & \mathbf{0}_{(m-r)\times(n-r)} \end{array}\right].$$

If $s > r$ then (i) U_{11} would have at least one row of $0_{\mathbb{F}}$'s meaning that U_{11} is non-invertible by Lemma 144. (ii) $U_{21} = \mathbf{0}_{(m-s)\times s}$. Thus from (i) and (ii) and from Lemma 146, U is not invertible, which is a contradiction. ❑

155 Corollary Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, $B \in \mathbf{M}_{n\times p}(\mathbb{F})$. If A is invertible then $\textbf{rank}\,(AB) = \textbf{rank}\,(B)$. If B is invertible then $\textbf{rank}\,(AB) = \textbf{rank}\,(A)$.

Proof: *Using Theorem 154, if A is invertible*

$$\textbf{rank}\,(AB) \leq \textbf{rank}\,(B) = \textbf{rank}\,\left(A^{-1}AB\right) \leq \textbf{rank}\,(AB),$$

and so $\textbf{rank}\,(B) = \textbf{rank}\,(AB)$. A similar argument works when B is invertible.

❑

156 Example Study the various possibilities for the rank of the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ b+c & c+a & a+b \\ bc & ca & ab \end{bmatrix}.$$

Solution: ► *Performing $R_2 - (b+c)R_1 \to R_2$ and $R_3 - bcR_1 \to R_3$, we find*

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & a-b & a-c \\ 0 & 0 & (b-c)(a-c) \end{bmatrix}.$$

Performing $C_2 - C_1 \to C_2$ and $C_3 - C_1 \to C_3$, we find

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & a-b & a-c \\ 0 & 0 & (b-c)(a-c) \end{bmatrix}.$$

We now examine the various ways of getting rows consisting only of 0's. If $a = b = c$, the last two rows are 0-rows and so $\textbf{rank}\,(A) = 1$. If exactly two of a, b, c are equal, the last row is a 0-row, but the middle one is not, and so $\textbf{rank}\,(A) = 2$ in this case. If none of a, b, c are equal, then the rank is clearly 3. ◄

Homework

Problem 2.7.1 On a symmetric matrix $A \in \mathbf{M}_{n\times n}(\mathbb{R})$ with $n \geq 3$,

$$R_3 - 3R_1 \to R_3$$

successively followed by

$$C_3 - 3C_1 \to C_3$$

are performed. Is the resulting matrix still symmetric?

Problem 2.7.2 Find the rank of

$$\begin{bmatrix} a+1 & a+2 & a+3 & a+4 & a+5 \\ a+2 & a+3 & a+4 & a+5 & a+6 \\ a+3 & a+4 & a+5 & a+6 & a+7 \\ a+4 & a+5 & a+6 & a+7 & a+8 \end{bmatrix} \in \mathbf{M}_{5\times 5}(\mathbb{R}).$$

Problem 2.7.3 Let A, B be arbitrary $n \times n$ matrices over $\mathbb{R}$. Prove or disprove! $\mathbf{rank}\,(AB) = \mathbf{rank}\,(BA)$.

Problem 2.7.4 Determine the rank of the matrix $\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 2 & 0 & 0 & 2 \end{bmatrix}$.

Problem 2.7.5 Suppose that the matrix $\begin{bmatrix} 4 & 2 \\ x^2 & x \end{bmatrix} \in \mathbf{M}_{2\times 2}(\mathbb{R})$ has rank 1. How many possible values can x assume?

Problem 2.7.6 Demonstrate that a non-zero $n \times n$ matrix A over a field $\mathbb{F}$ has rank 1 if and only if A can be factored as $A = XY$, where $X \in \mathbf{M}_{n\times 1}(\mathbb{F})$ and $Y \in \mathbf{M}_{1\times n}(\mathbb{F})$.

Problem 2.7.7 Study the various possibilities for the rank of the matrix

$$\begin{bmatrix} 1 & a & 1 & b \\ a & 1 & b & 1 \\ 1 & b & 1 & a \\ b & 1 & a & 1 \end{bmatrix}$$

when $(a, b) \in \mathbb{R}^2$.

Problem 2.7.8 Find the rank of $\begin{bmatrix} 1 & -1 & 0 & 1 \\ m & 1 & -1 & -1 \\ 1 & -m & 1 & 0 \\ 1 & -1 & m & 2 \end{bmatrix}$ as a function of $m \in \mathbb{C}$.

Problem 2.7.9 Determine the rank of the matrix $\begin{bmatrix} a^2 & ab & ab & b^2 \\ ab & a^2 & b^2 & ab \\ ab & b^2 & a^2 & ab \\ b^2 & ab & ab & a^2 \end{bmatrix}$.

Problem 2.7.10 Determine the rank of the matrix $\begin{bmatrix} 1 & 1 & 1 & 1 \\ a & b & a & b \\ c & c & d & d \\ ac & bc & ad & bd \end{bmatrix}$.

Problem 2.7.11 Let $A \in \mathbf{M}_{3\times 2}(\mathbb{R})$, $B \in \mathbf{M}_{2\times 2}(\mathbb{R})$, and $C \in \mathbf{M}_{2\times 3}(\mathbb{R})$ be such that $ABC = \begin{bmatrix} 1 & 1 & 2 \\ -2 & x & 1 \\ 1 & -2 & 1 \end{bmatrix}$. Find x.

Problem 2.7.12 Let B be the matrix obtained by adjoining a row (or column) to a matrix A. Prove that either $\mathbf{rank}(B) = \mathbf{rank}(A)$ or $\mathbf{rank}(B) = \mathbf{rank}(A) + 1$.

Problem 2.7.13 Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$. Prove that $\mathbf{rank}(A) = \mathbf{rank}(AA^T)$. Find a counterexample in the case $A \in \mathbf{M}_{n\times n}(\mathbb{C})$.

Problem 2.7.14 Prove that the rank of a skew-symmetric matrix with real number entries is an even number.

2.8 Rank and Invertibility

157 Theorem A matrix $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ is left-invertible if and only if $\mathbf{rank}(A) = n$. A matrix $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ is right-invertible if and only if $\mathbf{rank}(A) = m$.

Proof: *Observe that we always have* $\mathbf{rank}(A) \leq n$. *If* A *is left invertible, then* $\exists L \in \mathbf{M}_{n\times m}(\mathbb{F})$ *such that* $LA = \mathbf{I}_n$. *By Theorem 154,*

$$n = \mathbf{rank}(\mathbf{I}_n) = \mathbf{rank}(LA) \leq \mathbf{rank}(A),$$

whence the two inequalities give $\mathbf{rank}(A) = n$.

Conversely, assume that $\mathbf{rank}(A) = n$. *Then* $\mathbf{rank}(A^T) = n$ *by Corollary 150, and so by*

Theorem 149 there exist $P \in \mathbf{GL}_m(\mathbb{F})$, $Q \in \mathbf{GL}_n(\mathbb{F})$, *such that*

$$PAQ = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0}_{(m-n)\times n} \end{bmatrix}, \quad Q^TA^TP^T = \begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{n\times(m-n)} \end{bmatrix}.$$

This gives

$$\begin{aligned} Q^TA^TP^TPAQ = \mathbf{I}_n &\implies A^TP^TPA = (Q^T)^{-1}Q^{-1} \\ &\implies ((Q^T)^{-1}Q^{-1})^{-1}A^TP^TPA = \mathbf{I}_n, \end{aligned}$$

and so $((Q^T)^{-1}Q^{-1})^{-1}A^TP^TP$ *is a left inverse for* A.

The right-invertibility case is argued similarly. ❑

By combining Theorem 157 and Theorem 124, the following corollary is thus immediate.

158 Corollary If $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ possesses a left inverse L and a right inverse R then $m = n$ and $L = R$.

We use Gauß-Jordan Reduction to find the inverse of $A \in \mathbf{GL}_n(\mathbb{F})$. We form the *augmented matrix* $T = [A|\mathbf{I}_n]$ which is obtained by putting A side by side with the identity matrix $\mathbf{I}_n$. We perform permissible row operations on T until instead of A we obtain $\mathbf{I}_n$, which will appear if the matrix is invertible. The matrix on the right will be A^{-1}. We finish with $[\mathbf{I}_n|A^{-1}]$.

☞ *If* $A \in \mathbf{M}_{n\times n}(\mathbb{R})$ *is non-invertible, then the left hand side in the procedure above will not reduce to* $\mathbf{I}_n$.

159 Example Find the inverse of the matrix $B \in \mathbf{M}_{3\times 3}(\mathbb{Z}_7)$,

$$B = \begin{bmatrix} \overline{6} & \overline{0} & \overline{1} \\ \overline{3} & \overline{2} & \overline{0} \\ \overline{1} & \overline{0} & \overline{1} \end{bmatrix}.$$

Solution: ▶ *We have*

$$\left[\begin{array}{ccc|ccc} \overline{6} & \overline{0} & \overline{1} & \overline{1} & \overline{0} & \overline{0} \\ \overline{3} & \overline{2} & \overline{0} & \overline{0} & \overline{1} & \overline{0} \\ \overline{1} & \overline{0} & \overline{1} & \overline{0} & \overline{0} & \overline{1} \end{array}\right] \overset{R_1 \leftrightarrow R_3}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} \overline{1} & \overline{0} & \overline{1} & \overline{0} & \overline{0} & \overline{1} \\ \overline{3} & \overline{2} & \overline{0} & \overline{0} & \overline{1} & \overline{0} \\ \overline{6} & \overline{0} & \overline{1} & \overline{1} & \overline{0} & \overline{0} \end{array}\right]$$

$$\underset{R_2 - \overline{3}R_1 \to R_2}{\overset{R_3 - \overline{6}R_1 \to R_3}{\rightsquigarrow}} \left[\begin{array}{ccc|ccc} \overline{1} & \overline{0} & \overline{1} & \overline{0} & \overline{0} & \overline{1} \\ \overline{0} & \overline{2} & \overline{4} & \overline{0} & \overline{1} & \overline{4} \\ \overline{0} & \overline{0} & \overline{2} & \overline{1} & \overline{0} & \overline{1} \end{array}\right]$$

$$\underset{5R_1 + R_3 \to R_1}{\overset{R_2 - \overline{2}R_3 \to R_2}{\rightsquigarrow}} \left[\begin{array}{ccc|ccc} \overline{5} & \overline{0} & \overline{0} & \overline{1} & \overline{0} & \overline{6} \\ \overline{0} & \overline{2} & \overline{0} & \overline{5} & \overline{1} & \overline{2} \\ \overline{0} & \overline{0} & \overline{2} & \overline{1} & \overline{0} & \overline{1} \end{array}\right]$$

$$\underset{\overline{4}R_2 \to R_2}{\overset{\overline{3}R_1 \to R_1;\ \overline{4}R_3 \to R_3}{\rightsquigarrow}} \left[\begin{array}{ccc|ccc} \overline{1} & \overline{0} & \overline{0} & \overline{3} & \overline{0} & \overline{4} \\ \overline{0} & \overline{1} & \overline{0} & \overline{6} & \overline{4} & \overline{1} \\ \overline{0} & \overline{0} & \overline{1} & \overline{4} & \overline{0} & \overline{4} \end{array}\right].$$

We conclude that

$$\begin{bmatrix} \overline{6} & \overline{0} & \overline{1} \\ \overline{3} & \overline{2} & \overline{0} \\ \overline{1} & \overline{0} & \overline{1} \end{bmatrix}^{-1} = \begin{bmatrix} \overline{3} & \overline{0} & \overline{4} \\ \overline{6} & \overline{4} & \overline{1} \\ \overline{4} & \overline{0} & \overline{4} \end{bmatrix}.$$

◀

160 Example Use Gauß-Jordan reduction to find the inverse of the matrix $A = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix}$. Also, find A^{2001}.

Solution: ▶ *Operating on the augmented matrix*

$$\left[\begin{array}{ccc|ccc} 0 & 1 & -1 & 1 & 0 & 0 \\ 4 & -3 & 4 & 0 & 1 & 0 \\ 3 & -3 & 4 & 0 & 0 & 1 \end{array}\right] \overset{R_2-R_3\to R_2}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 0 & 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 3 & -3 & 4 & 0 & 0 & 1 \end{array}\right]$$

$$\overset{R_3-3R_2\to R_3}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 0 & 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & -3 & 4 & 0 & -3 & 4 \end{array}\right]$$

$$\overset{R_3+3R_1\to R_3}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 0 & 1 & -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 3 & -3 & 4 \end{array}\right]$$

$$\overset{R_1+R_3\to R_1}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 0 & 1 & 0 & 4 & -3 & 4 \\ 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 3 & -3 & 4 \end{array}\right]$$

$$\overset{R_1\leftrightarrow R_2}{\rightsquigarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & -1 \\ 0 & 1 & 0 & 4 & -3 & 4 \\ 0 & 0 & 1 & 3 & -3 & 4 \end{array}\right].$$

Thus we deduce that

$$A^{-1} = \begin{bmatrix} 0 & 1 & -1 \\ 4 & -3 & 4 \\ 3 & -3 & 4 \end{bmatrix} = A.$$

From $A^{-1} = A$ *we deduce* $A^2 = \mathbf{I}_n$. *Hence* $A^{2000} = (A^2)^{1000} = \mathbf{I}_n^{1000} = \mathbf{I}_n$ *and* $A^{2001} = A(A^{2000}) = A\mathbf{I}_n = A$. ◀

161 Example Find the inverse of the triangular matrix $A \in \mathbf{M}_{n\times n}(\mathbb{R})$,

$$A = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Solution: ► *Form the augmented matrix*

$$\left[\begin{array}{ccccc|ccccc} 1 & 1 & 1 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 1 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 1 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 1 \end{array}\right],$$

and perform $R_k - R_{k+1} \to R_k$ *successively for* $k = 1, 2, \ldots, n-1$, *obtaining*

$$\left[\begin{array}{ccccc|ccccc} 1 & 0 & 0 & \cdots & 0 & 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 1 \end{array}\right],$$

whence

$$A^{-1} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

that is, the inverse of A *has* 1*'s on the diagonal and* -1*'s on the superdiagonal.* ◄

162 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ be a triangular matrix such that $a_{11}a_{22}\cdots a_{nn} \neq 0_{\mathbb{F}}$. Then A is invertible.

Proof: *Since the entry* $a_{kk} \neq 0_{\mathbb{F}}$ *we multiply the* k*-th row by* a_{kk}^{-1} *and then proceed to subtract the appropriate multiples of the preceding* $k-1$ *rows at each stage.* ❑

163 Example (Putnam Exam, 1969) Let A and B be matrices of size 3×2 and 2×3 respectively. Suppose that their product AB is given by

$$AB = \begin{bmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{bmatrix}.$$

Demonstrate that the product **BA** is given by

$$BA = \begin{bmatrix} 9 & 0 \\ 0 & 9 \end{bmatrix}.$$

Solution: ▶ *Observe that*

$$(AB)^2 = \begin{bmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{bmatrix}\begin{bmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{bmatrix} = \begin{bmatrix} 72 & 18 & -18 \\ 18 & 45 & 36 \\ -18 & 36 & 45 \end{bmatrix} = 9AB.$$

Performing $R_3 + R_2 \to R_3$, $R_1 - 4R_2 \to R_2$, *and* $2R_3 + R_1 \to R_3$ *in succession we see that*

$$\begin{bmatrix} 8 & 2 & -2 \\ 2 & 5 & 4 \\ -2 & 4 & 5 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & -18 & -18 \\ 2 & 5 & 4 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & -18 & 0 \\ 2 & 5 & -1 \\ 0 & 0 & 0 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 0 & -18 & 0 \\ 0 & 5 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

and so $\mathbf{rank}\,(AB) = 2$. *This entails that* $\mathbf{rank}\left((AB)^2\right) = 2$. *Now, since* BA *is a* 2×2 *matrix,* $\mathbf{rank}\,(BA) \leq 2$. *Also*

$$2 = \mathbf{rank}\left((AB)^2\right) = \mathbf{rank}\,(ABAB) \leq \mathbf{rank}\,(ABA) \leq \mathbf{rank}\,(BA),$$

and we must conclude that $\mathbf{rank}\,(BA) = 2$. *This means that* BA *is invertible and so*

$$\begin{aligned} (AB)^2 = 9AB \quad & \implies \quad A(BA - 9\mathbf{I}_2)B = \mathbf{0}_3 \\ & \implies \quad BA(BA - 9\mathbf{I}_2)BA = B\mathbf{0}_3A \\ & \implies \quad BA(BA - 9\mathbf{I}_2)BA = \mathbf{0}_2 \\ & \implies \quad (BA)^{-1}BA(BA - 9\mathbf{I}_2)BA(BA)^{-1} = (BA)^{-1}\mathbf{0}_2(BA)^{-1} \\ & \implies \quad BA - 9\mathbf{I}_2 = \mathbf{0}_2 \end{aligned}$$

◀

Homework

Problem 2.8.1 Find the inverse of the matrix

$$\begin{bmatrix} \overline{1} & \overline{2} & \overline{3} \\ \overline{2} & \overline{3} & \overline{1} \\ \overline{3} & \overline{1} & \overline{2} \end{bmatrix} \in \mathbf{M}_{3\times 3}(\mathbb{Z}_7).$$

Problem 2.8.2 Let $(A, B) \in \mathbf{M}_{3\times 3}(\mathbb{R})$ be given by

$$A = \begin{bmatrix} a & b & c \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & a \\ -1 & a & b \end{bmatrix}.$$

Find B^{-1} and prove that $A^T = BAB^{-1}$.

Problem 2.8.3 Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & x \end{bmatrix}$ where $x \neq 0$ is a real number. Find A^{-1}.

Problem 2.8.4 If the inverse of the matrix $M = \begin{bmatrix} 1 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ is the matrix $M^{-1} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & -1 & a \\ 1 & 1 & b \end{bmatrix}$, find (a, b).

Problem 2.8.5 Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$ and let $n > 0$ be an integer. Find $(A^n)^{-1}$.

Problem 2.8.6 Give an example of a 2×2 invertible matrix A over $\mathbb{R}$ such that $A + A^{-1}$ is the zero matrix.

Problem 2.8.7 Find all the values of the parameter a for which the matrix B given below is not invertible.

$$B = \begin{bmatrix} -1 & a+2 & 2 \\ 0 & a & 1 \\ 2 & 1 & a \end{bmatrix}$$

Problem 2.8.8 Find the inverse of the triangular matrix

$$\begin{bmatrix} a & 2a & 3a \\ 0 & b & 2b \\ 0 & 0 & c \end{bmatrix} \in \mathbf{M}_{3\times 3}(\mathbb{R})$$

assuming that $abc \neq 0$.

Problem 2.8.9 Under what conditions is the matrix

$$\begin{bmatrix} b & a & 0 \\ c & 0 & a \\ 0 & c & b \end{bmatrix}$$

invertible? Find the inverse under these conditions.

Problem 2.8.10 Let A and B be $n \times n$ matrices over a field $\mathbb{F}$ such that AB is invertible. Prove that both A and B must be invertible.

Problem 2.8.11 Find the inverse of the matrix

$$\begin{bmatrix} 1+a & 1 & 1 \\ 1 & 1+b & 1 \\ 1 & 1 & 1+c \end{bmatrix}$$

Problem 2.8.12 Prove that for the $n \times n$ $(n > 1)$ matrix

$$\begin{bmatrix} 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & \dots & 1 \\ 1 & 1 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 1 & \dots & 0 \end{bmatrix}^{-1} = \frac{1}{n-1}\begin{bmatrix} 2-n & 1 & 1 & \dots & 1 \\ 1 & 2-n & 1 & \dots & 1 \\ 1 & 1 & 2-n & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 1 & \dots & 2-n \end{bmatrix}$$

Problem 2.8.13 Prove that the $n \times n$ $(n > 1)$ matrix

$$\begin{bmatrix} 1+a & 1 & 1 & \dots & 1 \\ 1 & 1+a & 1 & \dots & 1 \\ 1 & 1 & 1+a & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 1 & \dots & 1+a \end{bmatrix}$$

has inverse

$$-\frac{1}{a(n+a)}\begin{bmatrix} 1-n-a & 1 & 1 & \dots & 1 \\ 1 & 1-n-a & 1 & \dots & 1 \\ 1 & 1 & 1-n-a & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & 1 & 1 & \dots & 1-n-a \end{bmatrix}$$

Problem 2.8.14 Prove that

$$\begin{bmatrix} 1 & 3 & 5 & 7 & \cdots & (2n-1) \\ (2n-1) & 1 & 3 & 5 & \cdots & (2n-3) \\ (2n-3) & (2n-1) & 1 & 3 & \cdots & (2n-5) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 3 & 5 & 7 & 9 & \cdots & 1 \end{bmatrix}$$

has inverse

$$\frac{1}{2n^3}\begin{bmatrix} 2-n^2 & 2+n^2 & 2 & 2 & \cdots & 2 \\ 2 & 2-n^2 & 2+n^2 & 2 & \cdots & 2 \\ 2 & 2 & 2-n^2 & 2+n^2 & \cdots & 2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2+n^2 & 2 & 2 & 2 & \cdots & 2-n^2 \end{bmatrix}.$$

Problem 2.8.15 Prove that the $n \times n$ $(n > 1)$ matrix

$$\begin{bmatrix} 1+a_1 & 1 & 1 & \ldots & 1 \\ 1 & 1+a_2 & 1 & \ldots & 1 \\ 1 & 1 & 1+a_3 & \ldots & 1 \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ 1 & 1 & 1 & \ldots & 1+a_n \end{bmatrix}$$

has inverse

$$-\frac{1}{s}\begin{bmatrix} \frac{1-a_1 s}{a_1^2} & \frac{1}{a_1 a_2} & \frac{1}{a_1 a_3} & \ldots & \frac{1}{a_1 a_n} \\ \frac{1}{a_2 a_1} & \frac{1-a_2 s}{a_2^2} & \frac{1}{a_2 a_3} & \ldots & \frac{1}{a_2 a_n} \\ \frac{1}{a_3 a_1} & \frac{1}{a_3 a_2} & \frac{1-a_3 s}{a_3^2} & \ldots & \frac{1}{a_3 a_1 n} \\ \vdots & \vdots & \vdots & \ldots & \vdots \\ \frac{1}{a_n a_1} & \frac{1}{a_n a_2} & \frac{1}{a_n a_3} & \ldots & \frac{1-a_n s}{a_n^2} \end{bmatrix},$$

where $s = 1 + \frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}$.

Problem 2.8.16 Let $A \in \mathbf{M}_{5\times 5}(\mathbb{R})$. Shew that if $\mathbf{rank}\left(A^2\right) < 5$, then $\mathbf{rank}\,(A) < 5$.

Problem 2.8.17 Let p be an odd prime. How many invertible 2×2 matrices are there with entries all in $\mathbb{Z}_p$?

Problem 2.8.18 Let A, B be matrices of the same size. Prove that $\mathbf{rank}\,(A + B) \leq \mathbf{rank}\,(A) + \mathbf{rank}\,(B)$.

Problem 2.8.19 Let $A \in \mathbf{M}_{3,2}(\mathbb{R})$ and $B \in \mathbf{M}_{2,3}(\mathbb{R})$ be matrices such that $AB = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}$. Prove that $BA = \mathbf{I}_2$.

Chapter 3

Linear Equations

3.1 Definitions

We can write a system of m linear equations in n variables over a field $\mathbb{F}$

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= y_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= y_2, \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= y_m, \end{aligned}$$

in matrix form as

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}. \tag{3.1}$$

We write the above matrix relation in the abbreviated form

$$AX = Y, \tag{3.2}$$

where A is the matrix of coefficients, X is the matrix of variables and Y is the matrix of constants. Most often we will dispense with the matrix of variables X and will simply write the *augmented matrix* of the system as

$$[A|Y] = \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & y_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & y_m \end{array}\right]. \tag{3.3}$$

164 Definition Let $AX = Y$ be as in 3.1. If $Y = \mathbf{0}_{m\times 1}$, then the system is called *homogeneous*, otherwise it is called *inhomogeneous*. The set

$$\{X \in \mathbf{M}_{n\times 1}(\mathbb{F}) : AX = \mathbf{0}_{m\times 1}\}$$

is called the *kernel* or *nullspace* of A and it is denoted by $\mathbf{ker}\,(A)$.

☞ *Observe that we always have* $\mathbf{0}_{n\times 1} \in \mathbf{ker}(A) \in \mathbf{M}_{m\times n}(\mathbb{F})$.

165 Definition A system of linear equations is *consistent* if it has a solution. If the system does not have a solution then we say that it is *inconsistent*.

166 Definition If a row of a matrix is non-zero, we call the first non-zero entry of this row a *pivot* for this row.

167 Definition A matrix $M \in \mathbf{M}_{m\times n}(\mathbb{F})$ is a *row-echelon* matrix if

- All the zero rows of M, if any, are at the bottom of M.
- For any two consecutive rows R_i and R_{i+1}, either R_{i+1} is all $0_{\mathbb{F}}$'s or the pivot of R_{i+1} is immediately to the right of the pivot of R_i.

The variables accompanying these pivots are called the *leading variables*. Those variables which are not leading variables are the *free parameters*.

168 Example The matrices

$$\begin{bmatrix} \textcircled{1} & 0 & 1 & 1 \\ 0 & 0 & \textcircled{2} & 2 \\ 0 & 0 & 0 & \textcircled{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} \textcircled{1} & 0 & 1 & 1 \\ 0 & 0 & 0 & \textcircled{1} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

are in row-echelon form, with the pivots circled, but the matrices

$$\begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

are not in row-echelon form.

☞ *Observe that given a matrix* $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, *by following Gauß-Jordan reduction à la Theorem 149, we can find a matrix* $P \in \mathbf{GL}_m(\mathbb{F})$ *such that* $PA = B$ *is in row-echelon form.*

169 Example Solve the system of linear equations

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 4 \\ -6 \end{bmatrix}.$$

Solution: ▶ *Observe that the matrix of coefficients is already in row-echelon form. Clearly every variable is a leading variable, and by back substitution*

$$2w = -6 \implies w = -\frac{6}{2} = -3,$$

$$z - w = 4 \implies z = 4 + w = 4 - 3 = 1,$$

$$2y + z = -1 \implies y = -\frac{1}{2} - \frac{1}{2}z = -1,$$

$$x + y + z + w = -3 \implies x = -3 - y - z - w = 0.$$

The (unique) solution is thus

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 1 \\ -3 \end{bmatrix}.$$

◀

170 Example Solve the system of linear equations

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \\ 4 \end{bmatrix}.$$

Solution: ▶ *The system is already in row-echelon form, and we see that x, y, z are leading variables while w is a free parameter. We put $w = t$. Using back substitution, and operating from the bottom up, we find*

$$z - w = 4 \implies z = 4 + w = 4 + t,$$

$$2y + z = -1 \implies y = -\frac{1}{2} - \frac{1}{2}z = -\frac{1}{2} - 2 - \frac{1}{2}t = -\frac{5}{2} - \frac{1}{2}t,$$

$$x + y + z + w = -3 \implies x = -3 - y - z - w = -3 + \frac{5}{2} + \frac{1}{2}t - 4 - t - t = -\frac{9}{2} - \frac{3}{2}t.$$

The solution is thus

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -\frac{9}{2} - \frac{3}{2}t \\ -\frac{5}{2} - \frac{1}{2}t \\ 4 + t \\ t \end{bmatrix}, \quad t \in \mathbb{R}.$$

◀

171 Example Solve the system of linear equations

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -3 \\ -1 \end{bmatrix}.$$

Solution: ► *We see that* x, y *are leading variables, while* z, w *are free parameters. We put* $z = s, w = t$*. Operating from the bottom up, we find*

$$2y + z = -1 \implies y = -\frac{1}{2} - \frac{1}{2}z = -\frac{1}{2} - \frac{1}{2}s,$$

$$x + y + z + w = -3 \implies x = -3 - y - z - w = -\frac{5}{2} - \frac{3}{2}s - t.$$

The solution is thus

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -\frac{5}{2} - \frac{3}{2}s - t \\ -\frac{1}{2} - \frac{1}{2}s \\ s \\ t \end{bmatrix}, \quad (s, t) \in \mathbb{R}^2.$$

◄

172 Example Find all the solutions of the system

$$x + \overline{2}y + \overline{2}z = \overline{0},$$

$$y + \overline{2}z = \overline{1},$$

working in $\mathbb{Z}_3$.

Solution: ► *The augmented matrix of the system is*

$$\left[\begin{array}{ccc|c} \overline{1} & \overline{2} & \overline{2} & \overline{0} \\ \overline{0} & \overline{1} & \overline{2} & \overline{1} \end{array}\right].$$

The system is already in row-echelon form and x, y *are leading variables while* z *is a free parameter. We find*

$$y = \overline{1} - \overline{2}z = \overline{1} + \overline{1}z,$$

and

$$x = -\overline{2}y - \overline{2}z = \overline{1} + \overline{2}z.$$

Thus

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{1} + \overline{2}z \\ \overline{1} + \overline{1}z \\ z \end{bmatrix}, \quad z \in \mathbb{Z}_3.$$

Letting $z = \overline{0}, \overline{1}, \overline{2}$ successively, we find the three solutions

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{bmatrix},$$

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{0} \\ \overline{2} \\ \overline{1} \end{bmatrix},$$

and

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{2} \\ \overline{0} \\ \overline{2} \end{bmatrix}.$$

◀

Homework

Problem 3.1.1 Find all the solutions in $\mathbb{Z}_3$ of the system

$$x + y + z + w = \overline{0},$$

$$\overline{2}y + w = \overline{2}.$$

Problem 3.1.2 In $\mathbb{Z}_7$, given that

$$\begin{bmatrix} \overline{1} & \overline{2} & \overline{3} \\ \overline{2} & \overline{3} & \overline{1} \\ \overline{3} & \overline{1} & \overline{2} \end{bmatrix}^{-1} = \begin{bmatrix} \overline{4} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{4} \\ \overline{0} & \overline{4} & \overline{2} \end{bmatrix},$$

find all solutions of the system

$$\overline{1}x + \overline{2}y + \overline{3}z = \overline{5};$$

$$\overline{2}x + \overline{3}y + \overline{1}z = \overline{6};$$

$$\overline{3}x + \overline{1}y + \overline{2}z = \overline{0}.$$

Problem 3.1.3 Solve in $\mathbb{Z}_{13}$:

$$x - \overline{2}y + z = \overline{5}, \qquad \overline{2}x + \overline{2}y = \overline{7}, \qquad \overline{5}x - \overline{3}y + \overline{4}z = \overline{1}.$$

Problem 3.1.4 Find, with proof, a polynomial $p(x)$ with real number coefficients and degree 3 such that

$$p(-1) = -10, \quad p(0) = -1, \quad p(1) = 2, \quad p(2) = 23.$$

Problem 3.1.5 This problem introduces Hill block ciphers, which are a way of encoding information with an *encoding matrix* $A \in \mathbf{M}_{n\times n}(\mathbb{Z}_{26})$, where n is a strictly positive integer. Split a plaintext into blocks of n letters, creating a series of $n \times 1$ matrices P_k, and consider the numerical equivalent ($A = 0$, $B = 1$, $C = 2$, ..., $Z = 25$) of each letter. The encoded message is the translation to letters of the $n \times 1$ matrices $C_k = AP_k \mod 26$.

For example, suppose you want to encode the message **COMMUNISTS EAT OFFAL** with the encoding matrix

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix},$$

a 3×3 matrix. First, split the plaintext into groups of three letters:

COM MUN IST SEA TOF FAL.

Form 3×1 matrices with each set of letters and find their numerical equivalent, for example,

$$P_1 = \begin{bmatrix} \mathbf{C} \\ \mathbf{O} \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} 2 \\ 14 \\ 12 \end{bmatrix}.$$

Find the product AP_1 modulo 26, and translate into letters:

$$AP_1 = \begin{bmatrix} 0 & 1 & 0 \\ 3 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 14 \\ 12 \end{bmatrix} = \begin{bmatrix} 14 \\ 6 \\ 24 \end{bmatrix} = \begin{bmatrix} \mathbf{O} \\ \mathbf{G} \\ \mathbf{Y} \end{bmatrix},$$

hence **COM** is encoded into **OGY**. Your task is to complete the encoding of the message.

Problem 3.1.6 Find all solutions in $\mathbb{Z}_{103}$, if any, to the system

$$\begin{array}{rcl} x_0 + x_1 & = & 0, \\ x_0 + x_2 & = & 1, \\ x_0 + x_3 & = & 2, \\ \vdots & \vdots & \vdots \\ x_0 + x_{100} & = & 99, \\ x_0 + x_1 + x_2 + \cdots + x_{100} & = & 4949. \end{array}$$

Hints: $0 + 1 + 2 + \cdots + 99 = 4950$, $\quad 99 \cdot 77 - 103 \cdot 74 = 1$.

3.2 Existence of Solutions

We now answer the question of deciding when a system of linear equations is solvable.

173 Lemma Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ be in row-echelon form, and let $X \in \mathbf{M}_{n\times 1}(\mathbb{F})$ be a matrix of variables. The homogeneous system $AX = \mathbf{0}_{m\times 1}$ of m linear equations in n variables has (i) a unique solution if $m = n$, (ii) multiple solutions if $m < n$.

Proof: *If $m = n$ then A is a square triangular matrix whose diagonal elements are different from $0_{\mathbb{F}}$. As such, it is invertible by virtue of Theorem 162. Thus*

$$AX = \mathbf{0}_{n\times 1} \implies X = A^{-1}\mathbf{0}_{n\times 1} = \mathbf{0}_{n\times 1}$$

so there is only the unique solution $X = \mathbf{0}_{n\times 1}$, called the trivial solution.

If $m < n$ then there are $n - m$ free variables. Letting these variables run through the elements of the field, we obtain multiple solutions. Thus if the field has infinitely many elements, we obtain infinitely many solutions, and if the field has k elements, we obtain k^{n-m} solutions. Observe that in this case there is always a non-trivial solution.

❑

174 Theorem Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, and let $X \in \mathbf{M}_{n\times 1}(\mathbb{F})$ be a matrix of variables. The homogeneous system $AX = \mathbf{0}_{m\times 1}$ of m linear equations in n variables always has a non-trivial solution if $m < n$.

Proof: *We can find a matrix $P \in \mathbf{GL}_m(\mathbb{F})$ such that $B = PA$ is in row-echelon form. Now*

$$AX = \mathbf{0}_{m\times 1} \iff PAX = \mathbf{0}_{m\times 1} \iff BX = \mathbf{0}_{m\times 1}.$$

That is, the systems $AX = \mathbf{0}_{m\times 1}$ and $BX = \mathbf{0}_{m\times 1}$ have the same set of solutions. But by Lemma 173 there is a non-trivial solution. ❑

175 Theorem (Kronecker-Capelli) Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, $Y \in \mathbf{M}_{m\times 1}(\mathbb{F})$ be constant matrices and $X \in \mathbf{M}_{n\times 1}(\mathbb{F})$ be a matrix of variables. The matrix equation $AX = Y$ is solvable if and only if

$$\mathbf{rank}\,(A) = \mathbf{rank}\,([A|Y]).$$

Proof: *Assume first that* $AX = Y$,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Let the columns of $[A|X]$ *be denoted by* $C_i, 1 \leq i \leq n$. *Observe that that* $[A|X] \in \mathbf{M}_{m\times(n+1)}(\mathbb{F})$ *and that the* $(n+1)$*-th column of* $[A|X]$ *is*

$$C_{n+1} = AX = \begin{bmatrix} x_1a_{11} + x_2a_{12} + \cdots + x_na_{1n} \\ x_1a_{21} + x_2a_{22} + \cdots + x_na_{2n} \\ \vdots \\ x_1a_{n1} + x_2a_{n2} + \cdots + x_na_{nn} \end{bmatrix} = \sum_{i=1}^{n} x_iC_i.$$

By performing $C_{n+1} - \sum_{j=1}^{n} x_jC_j \to C_{n+1}$ *on* $[A|Y] = [A|AX]$ *we obtain* $[A|0_{n\times1}]$. *Thus* **rank** $([A|Y]) =$ **rank** $([A|0_{n\times1}]) =$ **rank** (A).

Now assume that $r =$ **rank** $(A) =$ **rank** $([A|Y])$. *This means that adding an extra column to* A *does not change the rank, and hence, by a sequence column operations* $[A|Y]$ *is equivalent to* $[A|0_{n\times1}]$. *Observe that none of these operations is a permutation of the columns, since the first* n *columns of* $[A|Y]$ *and* $[A|0_{n\times1}]$ *are the same. This means that* Y *can be obtained from the columns* $C_i, 1 \leq i \leq n$ *of* A *by means of transvections and dilatations. But then*

$$Y = \sum_{i=1}^{n} x_iC_i.$$

The solutions sought is thus

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

❑

Problem 3.2.1 Let $A \in \mathbf{M}_{n\times p}(\mathbb{F})$, $B \in \mathbf{M}_{n\times q}(\mathbb{F})$ and put $C = [A \quad B] \in \mathbf{M}_{n\times(p+q)}(\mathbb{F})$ Prove that **rank** $(A) =$ **rank** $(C) \iff \exists P \in \mathbf{M}_p(q)$ such that $B = AP$.

3.3 Examples of Linear Systems

176 Example Use row reduction to solve the system

$$\begin{array}{ccccccccc} x & + & 2y & + & 3z & + & 4w & = & 8 \\ x & + & 2y & + & 4z & + & 7w & = & 12 \\ 2x & + & 4y & + & 6z & + & 8w & = & 16 \end{array}$$

Solution: ▶ *Form the expanded matrix of coefficients and apply row operations to obtain*

$$\left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 8 \\ 1 & 2 & 4 & 7 & 12 \\ 2 & 4 & 6 & 8 & 16 \end{array}\right] \overset{R_3-2R_1\to R_3}{\underset{R_2-R_1\to R_2}{\leadsto}} \left[\begin{array}{cccc|c} 1 & 2 & 3 & 4 & 8 \\ 0 & 0 & 1 & 3 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

The matrix is now in row-echelon form. The variables x *and* z *are the pivots, so* w *and* y *are free. Setting* $w = s, y = t$ *we have*

$$z = 4 - 3s,$$

$$x = 8 - 4w - 3z - 2y = 8 - 4s - 3(4 - 3s) - 2t = -4 + 5s - 2t.$$

Hence the solution is given by

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -4 + 5s - 2t \\ t \\ 4 - 3s \\ s \end{bmatrix}.$$

◀

177 Example Find $\alpha \in \mathbb{R}$ such that the system

$$x + y - z = 1,$$

$$2x + 3y + \alpha z = 3,$$

$$x + \alpha y + 3z = 2,$$

posses (i) no solution, (ii) infinitely many solutions, (iii) a unique solution.

Solution: ▶ *The augmented matrix of the system is*

$$\left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 2 & 3 & \alpha & 3 \\ 1 & \alpha & 3 & 2 \end{array}\right].$$

By performing $R_2 - 2R_1 \to R_2$ *and* $R_3 - R_1 \to R_3$ *we obtain*

$$\leadsto \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & \alpha + 2 & 1 \\ 0 & \alpha - 1 & 4 & 1 \end{array}\right].$$

By performing $R_3 - (\alpha - 1)R_2 \to R_3$ *on this last matrix we obtain*

$$\leadsto \left[\begin{array}{ccc|c} 1 & 1 & -1 & 1 \\ 0 & 1 & \alpha + 2 & 1 \\ 0 & 0 & (\alpha - 2)(\alpha + 3) & \alpha - 2 \end{array}\right].$$

If $\alpha = -3$, we obtain no solution. If $\alpha = 2$, there is an infinity of solutions

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5t \\ 1-4t \\ t \end{bmatrix}, \quad t \in \mathbb{R}.$$

If $\alpha \neq 2$ and $\alpha \neq 3$, there is a unique solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ \dfrac{1}{\alpha+3} \\ \dfrac{1}{\alpha+3} \end{bmatrix}.$$

◀

178 Example Solve the system

$$\begin{bmatrix} \overline{6} & \overline{0} & \overline{1} \\ \overline{3} & \overline{2} & \overline{0} \\ \overline{1} & \overline{0} & \overline{1} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{1} \\ \overline{0} \\ \overline{2} \end{bmatrix},$$

for $(x, y, z) \in (\mathbb{Z}_7)^3$.

Solution: ▶ *Performing operations on the augmented matrix we have*

$$\left[\begin{array}{ccc|c} \overline{6} & \overline{0} & \overline{1} & \overline{1} \\ \overline{3} & \overline{2} & \overline{0} & \overline{0} \\ \overline{1} & \overline{0} & \overline{1} & \overline{2} \end{array}\right] \overset{R_1 \leftrightarrow R_3}{\rightsquigarrow} \left[\begin{array}{ccc|c} \overline{1} & \overline{0} & \overline{1} & \overline{2} \\ \overline{3} & \overline{2} & \overline{0} & \overline{0} \\ \overline{6} & \overline{0} & \overline{1} & \overline{1} \end{array}\right]$$

$$\underset{R_2 - \overline{3}R_1 \to R_2}{\overset{R_3 - \overline{6}R_1 \to R_3}{\rightsquigarrow}} \left[\begin{array}{ccc|c} \overline{1} & \overline{0} & \overline{1} & \overline{2} \\ \overline{0} & \overline{2} & \overline{4} & \overline{1} \\ \overline{0} & \overline{0} & \overline{2} & \overline{3} \end{array}\right]$$

This gives

$$\overline{2}z = \overline{3} \implies z = \overline{5},$$

$$\overline{2}y = \overline{1} - \overline{4}z = \overline{2} \implies y = \overline{1},$$

$$x = \overline{2} - z = \overline{4}.$$

The solution is thus

$$(x, y, z) = (\overline{4}, \overline{1}, \overline{5}).$$

◀

Homework

Problem 3.3.1 Find the general solution to the system

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 2 & 1 & 2 & 1 & 2 \\ 4 & 2 & 4 & 2 & 4 \\ 1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ f \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

or shew that there is no solution.

Problem 3.3.2 Find all solutions of the system

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 2 \\ 1 & 1 & 1 & 3 & 3 \\ 1 & 1 & 4 & 4 & 4 \\ 1 & 2 & 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ f \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 7 \\ 6 \\ 9 \end{bmatrix},$$

if any.

Problem 3.3.3 Study the system

$$x + 2my + z = 4m;$$
$$2mx + y + z = 2;$$
$$x + y + 2mz = 2m^2,$$

with real parameter m. You must determine, with proof, for which m this system has (i) no solution, (ii) exactly one solution, and (iii) infinitely many solutions.

Problem 3.3.4 Study the following system of linear equations with parameter a.

$$(2a-1)x + ay - (a+1)z = 1,$$

$$ax + y - 2z = 1,$$

$$2x + (3-a)y + (2a-6)z = 1.$$

You must determine for which a there is: (i) no solution, (ii) a unique solution, (iii) infinitely many solutions.

Problem 3.3.5 Determine the values of the parameter m for which the system

$$\begin{array}{lclcll} x & + & y & + & (1-m)z & = & m+2 \\ (1+m)x & - & y & + & 2z & = & 0 \\ 2x & - & my & + & 3z & = & m+2 \end{array}$$

is solvable.

Problem 3.3.6 Determine the values of the parameter m for which the system

$$\begin{array}{rcrcrcrcl} x & + & y & + & z & + & t & = & 4a \\ x & - & y & - & z & + & t & = & 4b \\ -x & - & y & + & z & + & t & = & 4c \\ x & - & y & + & z & - & t & = & 4d \end{array}$$

is solvable.

Problem 3.3.7 It is known that the system

$$\begin{aligned} ay + bx &= c; \\ cx + az &= b; \\ bz + cy &= a \end{aligned}$$

possesses a unique solution. What conditions must $(a, b, c) \in \mathbb{R}^3$ fulfill in this case? Find this unique solution.

Problem 3.3.8 For which values of the real parameter a does the following system have (i) no solutions, (ii) exactly one solution, (iii) infinitely many solutions?

$$\begin{array}{rcrcrcr} (1-a)x & + & (2a+1)y & + & (2a+2)z & = & a, \\ ax & + & ay & & & = & 2a+2, \\ 2x & + & (a+1)y & + & (a-1)z & = & a^2-2a+9. \end{array}$$

Problem 3.3.9 Find strictly positive real numbers x, y, z such that

$$\begin{aligned} x^3y^2z^6 &= 1 \\ x^4y^5z^{12} &= 2 \\ x^2y^2z^5 &= 3. \end{aligned}$$

Problem 3.3.10 (Leningrad Mathematical Olympiad, 1987, Grade 5) The numbers $1, 2, \ldots, 16$ are arranged in a 4×4 matrix A as shewn below. We may add 1 to all the numbers of any row or subtract 1 from all numbers of any column. Using only the allowed operations, how can we obtain A^T?

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$$

Problem 3.3.11 (International Mathematics Olympiad, 1963) Find all solutions x_1, x_2, x_3, x_4, x_5 of the system

$$\begin{aligned} x_5 + x_2 &= yx_1; \\ x_1 + x_3 &= yx_2; \\ x_2 + x_4 &= yx_3; \\ x_3 + x_5 &= yx_4; \\ x_4 + x_1 &= yx_5, \end{aligned}$$

where y is a parameter.

Chapter 4

Vector Spaces

4.1 Vector Spaces

179 Definition A *vector space* $\langle V, +, \cdot, \mathbb{F}\rangle$ over a field $\langle \mathbb{F}, +, \rangle$ is a non-empty set V whose elements are called *vectors*, possessing two operations $+$ (vector addition), and $\cdot$ (scalar multiplication) which satisfy the following axioms.

$$\forall(\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c}) \in V^3, \ \forall(\alpha, \beta) \in \mathbb{F}^2,$$

VS1 **Closure under vector addition** :

$$\overrightarrow{a} + \overrightarrow{b} \in V, \tag{4.1}$$

VS2 **Closure under scalar multiplication**

$$\alpha\overrightarrow{a} \in V, \tag{4.2}$$

VS3 **Commutativity**

$$\overrightarrow{a} + \overrightarrow{b} = \overrightarrow{b} + \overrightarrow{a} \tag{4.3}$$

VS4 **Associativity**

$$(\overrightarrow{a} + \overrightarrow{b}) + \overrightarrow{c} = \overrightarrow{a} + (\overrightarrow{b} + \overrightarrow{c}) \tag{4.4}$$

VS5 **Existence of an additive identity**

$$\exists\ \overrightarrow{0} \in V : \overrightarrow{a} + \overrightarrow{0} = \overrightarrow{a} + \overrightarrow{0} = \overrightarrow{a} \tag{4.5}$$

VS6 **Existence of additive inverses**

$$\exists\ -\overrightarrow{a} \in V : \overrightarrow{a} + (-\overrightarrow{a}) = (-\overrightarrow{a}) + \overrightarrow{a} = \overrightarrow{0} \tag{4.6}$$

VS7 **Distributive Law**

$$\alpha(\overrightarrow{a} + \overrightarrow{b}) = \alpha\overrightarrow{a} + \alpha\overrightarrow{b} \tag{4.7}$$

VS8 **Distributive Law**

$$(\alpha + \beta)\overrightarrow{a} = \alpha\overrightarrow{a} + \beta\overrightarrow{a} \tag{4.8}$$

VS9

$$1_{\mathbb{F}}\overrightarrow{a} = \overrightarrow{a} \tag{4.9}$$

VS10

$$(\alpha\beta)\overrightarrow{a} = \alpha(\beta\overrightarrow{a}) \tag{4.10}$$

180 Example If n is a positive integer, then $\langle \mathbb{F}^n, +, \cdot, \mathbb{F}\rangle$ is a vector space by defining

$$(a_1, a_2, \ldots, a_n) + (b_1, b_2, \ldots, b_n) = (a_1 + b_1, a_2 + b_2, \ldots, a_n + b_n),$$

$$\lambda(a_1, a_2, \ldots, a_n) = (\lambda a_1, \lambda a_2, \ldots, \lambda a_n).$$

In particular, $\langle \mathbb{Z}_2^2, +, \cdot, \mathbb{Z}_2\rangle$ is a vector space with only four elements and we have seen the two-dimensional and tridimensional spaces $\langle \mathbb{R}^2, +, \cdot, \mathbb{R}\rangle$ and $\langle \mathbb{R}^3, +, \cdot, \mathbb{R}\rangle$.

181 Example $\langle \mathbf{M}_{m\times n}(\mathbb{F}), +, \cdot, \mathbb{F}\rangle$ is a vector space under matrix addition and scalar multiplication of matrices.

182 Example If

$$\mathbb{F}[x] = \{a_0 + a_1x + a_2x + \cdots + a_nx^n : a_i \in \mathbb{F}, \quad n \in \mathbb{N}\}$$

denotes the set of polynomials with coefficients in a field $\langle \mathbb{F}, +, \rangle$ then $\langle \mathbb{F}[x], +, \cdot, \mathbb{F}\rangle$ is a vector space, under polynomial addition and scalar multiplication of a polynomial.

183 Example If

$$\mathbb{F}_n[x] = \{a_0 + a_1x + a_2x + \cdots + a_kx^k : a_i \in \mathbb{F}, \quad n \in \mathbb{N}, k \leq n\}$$

denotes the set of polynomials with coefficients in a field $\langle \mathbb{F}, +, \rangle$ and degree at most n, then $\langle \mathbb{F}_n[x], +, \cdot, \mathbb{F}\rangle$ is a vector space, under polynomial addition and scalar multiplication of a polynomial.

184 Example Let $k \in \mathbb{N}$ and let $C^k(\mathbb{R}^{[a;b]})$ denote the set of k-fold continuously differentiable real-valued functions defined on the interval $[a; b]$. Then $C^k(\mathbb{R}^{[a;b]})$ is a vector space under addition of functions and multiplication of a function by a scalar.

185 Example Let $p \in]1; +\infty[$. Consider the set of sequences $\{a_n\}_{n=0}^{\infty}, \quad a_n \in \mathbb{C}$,

$$l^p = \left\{ \{a_n\}_{n=0}^{\infty} : \sum_{n=0}^{\infty} |a_n|^p < +\infty \right\}.$$

Then l^p is a vector space by defining addition as termwise addition of sequences and scalar multiplication as termwise multiplication:

$$\{a_n\}_{n=0}^{\infty} + \{b_n\}_{n=0}^{\infty} = \{(a_n + b_n)\}_{n=0}^{\infty},$$

$$\lambda\{a_n\}_{n=0}^{\infty} = \{\lambda a_n\}_{n=0}^{\infty}, \quad \lambda \in \mathbb{C}.$$

All the axioms of a vector space follow trivially from the fact that we are adding complex numbers, except that we must prove that in l^p there is closure under addition and scalar multiplication. Since $\sum_{n=0}^{\infty} |a_n|^p < +\infty \implies \sum_{n=0}^{\infty} |\lambda a_n|^p < +\infty$ closure under scalar multiplication follows easily. To prove closure under addition, observe that if $z \in \mathbb{C}$ then $|z| \in \mathbb{R}_+$ and so by the Minkowski Inequality Theorem 405 we have

$$\begin{aligned} \left(\sum_{n=0}^{N} |a_n + b_n|^p\right)^{1/p} &\leq \left(\sum_{n=0}^{N} |a_n|^p\right)^{1/p} + \left(\sum_{n=0}^{N} |b_n|^p\right)^{1/p} \\ &\leq \left(\sum_{n=0}^{\infty} |a_n|^p\right)^{1/p} + \left(\sum_{n=0}^{\infty} |b_n|^p\right)^{1/p}. \end{aligned} \tag{4.11}$$

This in turn implies that the series on the left in (4.11) converges, and so we may take the limit as $N \to +\infty$ obtaining

$$\left(\sum_{n=0}^{\infty} |a_n + b_n|^p\right)^{1/p} \leq \left(\sum_{n=0}^{\infty} |a_n|^p\right)^{1/p} + \left(\sum_{n=0}^{\infty} |b_n|^p\right)^{1/p}. \tag{4.12}$$

Now (4.12) implies that the sum of two sequences in l^p is also in l^p, which demonstrates closure under addition.

186 Example The set

$$\mathbf{V} = \{a + b\sqrt{2} + c\sqrt{3} : (a, b, c) \in \mathbb{Q}^3\}$$

with addition defined as

$$(a + b\sqrt{2} + c\sqrt{3}) + (a' + b'\sqrt{2} + c'\sqrt{3}) = (a + a') + (b + b')\sqrt{2} + (c + c')\sqrt{3},$$

and scalar multiplication defined as

$$\lambda(a + b\sqrt{2} + c\sqrt{3}) = (\lambda a) + (\lambda b)\sqrt{2} + (\lambda c)\sqrt{3},$$

constitutes a vector space over $\mathbb{Q}$.

187 Theorem In any vector space $\langle \mathbf{V}, +, \cdot, \mathbb{F}\rangle$,

$$\forall \ \alpha \in \mathbb{F}, \quad \alpha\overrightarrow{0} = \overrightarrow{0}.$$

Proof: *We have*

$$\alpha\overrightarrow{0} = \alpha(\overrightarrow{0} + \overrightarrow{0}) = \alpha\overrightarrow{0} + \alpha\overrightarrow{0}.$$

Hence

$$\alpha\overrightarrow{0} - \alpha\overrightarrow{0} = \alpha\overrightarrow{0},$$

or

$$\overrightarrow{0} = \alpha\overrightarrow{0},$$

proving the theorem. ❑

188 Theorem In any vector space $\langle \mathbf{V}, +, \cdot, \mathbb{F}\rangle$,

$$\forall \ \overrightarrow{v} \in \mathbf{V}, \quad 0_{\mathbb{F}}\overrightarrow{v} = \overrightarrow{0}.$$

Proof: *We have*

$$0_{\mathbb{F}}\overrightarrow{v} = (0_{\mathbb{F}} + 0_{\mathbb{F}})\overrightarrow{v} = 0_{\mathbb{F}}\overrightarrow{v} + 0_{\mathbb{F}}\overrightarrow{v}.$$

Therefore

$$0_{\mathbb{F}}\overrightarrow{v} - 0_{\mathbb{F}}\overrightarrow{v} = 0_{\mathbb{F}}\overrightarrow{v},$$

or

$$\overrightarrow{0} = 0_{\mathbb{F}}\overrightarrow{v},$$

proving the theorem. ❑

189 Theorem In any vector space $\langle \mathbf{V}, +, \cdot, \mathbb{F}\rangle$, $\alpha \in \mathbb{F}$, $\overrightarrow{v} \in \mathbf{V}$,

$$\alpha\overrightarrow{v} = \overrightarrow{0} \quad \implies \quad \alpha = 0_{\mathbb{F}} \quad \vee \quad \overrightarrow{v} = \overrightarrow{0}.$$

Proof: *Assume that $\alpha \neq 0_{\mathbb{F}}$. Then α possesses a multiplicative inverse α^{-1} such that $\alpha^{-1}\alpha = 1_{\mathbb{F}}$. Thus*

$$\alpha\overrightarrow{v} = \overrightarrow{0} \implies \alpha^{-1}\alpha\overrightarrow{v} = \alpha^{-1}\overrightarrow{0}.$$

By Theorem 188, $\alpha^{-1}\overrightarrow{0} = \overrightarrow{0}$. Hence

$$\alpha^{-1}\alpha\overrightarrow{v} = \overrightarrow{0}.$$

Since by Axiom 4.9, we have $\alpha^{-1}\alpha\overrightarrow{v} = 1_{\mathbb{F}}\overrightarrow{v} = \overrightarrow{v}$, and so we conclude that $\overrightarrow{v} = \overrightarrow{0}$. ❑

190 Theorem In any vector space $\langle \mathbf{V}, +, \cdot, \mathbb{F}\rangle$,

$$\forall\alpha \in \mathbb{F}, \quad \forall \ \overrightarrow{v} \in \mathbf{V}, \quad (-\alpha)\overrightarrow{v} = \alpha(-\overrightarrow{v}) = -(\alpha\overrightarrow{v}).$$

Proof: *We have*

$$0_{\mathbb{F}}\overrightarrow{v} = (\alpha + (-\alpha))\overrightarrow{v} = \alpha\overrightarrow{v} + (-\alpha)\overrightarrow{v},$$

whence

$$-(\alpha\overrightarrow{v}) + 0_{\mathbb{F}}\overrightarrow{v} = (-\alpha)\overrightarrow{v},$$

that is

$$-(\alpha\overrightarrow{v}) = (-\alpha)\overrightarrow{v}.$$

Similarly,

$$\overrightarrow{0} = \alpha(\overrightarrow{v} - \overrightarrow{v}) = \alpha\overrightarrow{v} + \alpha(-\overrightarrow{v}),$$

whence

$$-(\alpha\overrightarrow{v}) + \overrightarrow{0} = \alpha(-\overrightarrow{v}),$$

that is

$$-(\alpha\overrightarrow{v}) = \alpha(-\overrightarrow{v}),$$

proving the theorem. ❑

Homework

Problem 4.1.1 Is $\mathbb{R}^2$ with vector addition and scalar multiplication defined as

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \end{bmatrix}, \quad \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ 0 \end{bmatrix}$$

a vector space?

Problem 4.1.2 Demonstrate that the commutativity axiom 4.3 is redundant.

Problem 4.1.3 Let $V = \mathbb{R}^+ =]0; +\infty[$, the positive real numbers and $F = \mathbb{R}$, the real numbers. Demonstrate that V is a vector space over $\mathbb{F}$ if vector addition is defined as $a \oplus b = ab$, $(a, b) \in (\mathbb{R}^+)^2$ and scalar multiplication is defined as $\alpha \otimes a = a^\alpha$, $(\alpha, a) \in (\mathbb{R}, \mathbb{R}^+)$.

Problem 4.1.4 Let $\mathbb{C}$ denote the complex numbers and $\mathbb{R}$ denote the real numbers. Is $\mathbb{C}$ a vector space over $\mathbb{R}$ under ordinary addition and multiplication? Is $\mathbb{R}$ a vector space over $\mathbb{C}$?

Problem 4.1.5 Construct a vector space with exactly 8 elements.

Problem 4.1.6 Construct a vector space with exactly 9 elements.

4.2 Vector Subspaces

191 Definition Let $\langle V, +, \cdot, \mathbb{F}\rangle$ be a vector space. A non-empty subset $U \subseteq V$ which is also a vector space under the inherited operations of V is called a *vector subspace of* V.

192 Example Trivially, $X_1 = \{\overrightarrow{0}\}$ and $X_2 = V$ are vector subspaces of V.

193 Theorem Let $\langle V, +, \cdot, \mathbb{F}\rangle$ be a vector space. Then $U \subseteq V$, $U \neq \varnothing$ is a subspace of V if and only if $\forall\alpha \in \mathbb{F}$ and $\forall(\overrightarrow{a}, \overrightarrow{b}) \in U^2$ it is verified that

$$\overrightarrow{a} + \alpha\overrightarrow{b} \in U.$$

Proof: *Observe that U inherits commutativity, associativity and the distributive laws from V. Thus a non-empty $U \subseteq V$ is a vector subspace of V if (i) U is closed under scalar multiplication, that is, if $\alpha \in \mathbb{F}$ and $\overrightarrow{v} \in U$, then $\alpha\overrightarrow{v} \in U$; (ii) U is closed under vector addition, that is, if $(\overrightarrow{u}, \overrightarrow{v}) \in U^2$, then $\overrightarrow{u} + \overrightarrow{v} \in U$. Observe that (i) gives the existence of inverses in U, for take $\alpha = -1_{\mathbb{F}}$ and so $\overrightarrow{v} \in U \implies -\overrightarrow{v} \in U$. This coupled with (ii) gives the existence of the zero-vector, for $\overrightarrow{0} = \overrightarrow{v} - \overrightarrow{v} \in U$. Thus we need to prove that if a non-empty subset of V satisfies the property* *stated in the Theorem then it is closed under scalar multiplication and vector addition, and vice-versa, if a non-empty subset of* V *is closed under scalar multiplication and vector addition, then it satisfies the property stated in the Theorem. But this is trivial.* ❑

194 Example Shew that $X = \{A \in \mathbf{M}_{n\times n}(\mathbb{F}) : \mathbf{tr}(A) = 0_{\mathbb{F}}\}$ is a subspace of $\mathbf{M}_{n\times n}(\mathbb{F})$.

Solution: ▶ *Take* $A, B \in X$, $\alpha \in \mathbb{R}$. *Then*

$$\mathbf{tr}(A + \alpha B) = \mathbf{tr}(A) + \alpha\mathbf{tr}(B) = 0_{\mathbb{F}} + \alpha(0_{\mathbb{F}}) = 0_{\mathbb{F}}.$$

Hence $A + \alpha B \in X$, *meaning that* X *is a subspace of* $\mathbf{M}_{n\times n}(\mathbb{F})$. ◀

195 Example Let $U \in \mathbf{M}_{n\times n}(\mathbb{F})$ be an arbitrary but fixed. Shew that

$$\mathscr{C}_U = \{A \in \mathbf{M}_{n\times n}(\mathbb{F}) : AU = UA\}$$

is a subspace of $\mathbf{M}_{n\times n}(\mathbb{F})$.

Solution: ▶ *Take* $(A, B) \in (\mathscr{C}_U)^2$. *Then* $AU = UA$ *and* $BU = UB$. *Now*

$$(A + \alpha B)U = AU + \alpha BU = UA + \alpha UB = U(A + \alpha B),$$

meaning that $A + \alpha B \in \mathscr{C}_U$. *Hence* $\mathscr{C}_U$ *is a subspace of* $\mathbf{M}_{n\times n}(\mathbb{F})$. $\mathscr{C}_U$ *is called the* commutator *of* U. ◀

196 Theorem Let $X \subseteq V$, $Y \subseteq V$ be vector subspaces of a vector space $\langle V, +, \cdot, \mathbb{F}\rangle$. Then their intersection $X \cap Y$ is also a vector subspace of V.

Proof: *Let* $\alpha \in \mathbb{F}$ *and* $(\overrightarrow{a}, \overrightarrow{b}) \in (X \cap Y)^2$. *Then clearly* $(\overrightarrow{a}, \overrightarrow{b}) \in X$ *and* $(\overrightarrow{a}, \overrightarrow{b}) \in Y$. *Since* X *is a vector subspace,* $\overrightarrow{a} + \alpha\overrightarrow{b} \in X$ *and since* Y *is a vector subspace,* $\overrightarrow{a} + \alpha\overrightarrow{b} \in Y$. *Thus*

$$\overrightarrow{a} + \alpha\overrightarrow{b} \in X \cap Y$$

and so $X \cap Y$ *is a vector subspace of* V *by virtue of Theorem 193.* ❑

☞ *We we will soon see that the only vector subspaces of* $\langle \mathbb{R}^2, +, \cdot, \mathbb{R}\rangle$ *are the set containing the zero-vector, any line through the origin, and* $\mathbb{R}^2$ *itself. The only vector subspaces of* $\langle \mathbb{R}^3, +, \cdot, \mathbb{R}\rangle$ *are the set containing the zero-vector, any line through the origin, any plane containing the origin and* $\mathbb{R}^3$ *itself.*

Homework

Problem 4.2.1 Prove that

$$X = \left\{ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathbb{R}^4 : a - b - 3d = 0 \right\}$$

is a vector subspace of $\mathbb{R}^4$.

Problem 4.2.2 Prove that

$$X = \left\{ \begin{bmatrix} a \\ 2a - 3b \\ 5b \\ a + 2b \\ a \end{bmatrix} : a, b \in \mathbb{R} \right\}$$

is a vector subspace of $\mathbb{R}^5$.

Problem 4.2.3 Let $A \in \mathbf{M}_{m\times n}(\mathbb{F})$ be a fixed matrix. Demonstrate that

$$S = \{X \in \mathbf{M}_{n\times 1}(\mathbb{F}) : AX = \mathbf{0}_{m\times 1}\}$$

is a subspace of $\mathbf{M}_{n\times 1}(\mathbb{F})$.

Problem 4.2.4 Prove that the set $X \subseteq \mathbf{M}_{n\times n}(\mathbb{F})$ of upper triangular matrices is a subspace of $\mathbf{M}_{n\times n}(\mathbb{F})$.

Problem 4.2.5 Prove that the set $X \subseteq \mathbf{M}_{n\times n}(\mathbb{F})$ of symmetric matrices is a subspace of $\mathbf{M}_{n\times n}(\mathbb{F})$.

Problem 4.2.6 Prove that the set $X \subseteq \mathbf{M}_{n\times n}(\mathbb{F})$ of skew-symmetric matrices is a subspace of $\mathbf{M}_{n\times n}(\mathbb{F})$.

Problem 4.2.7 Prove that the following subsets are **not** subspaces of the given vector space. Here you must say which of the axioms for a vector space fail.

❶ $$\left\{\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a, b \in \mathbb{R}, a^2 + b^2 = 1\right\} \subseteq \mathbb{R}^3$$

❷ $$\left\{\begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a, b \in \mathbb{R}^2, ab = 0\right\} \subseteq \mathbb{R}^3$$

❸ $$\left\{\begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} : (a, b) \in \mathbb{R}^2, a + b^2 = 0\right\} \subseteq \mathbf{M}_{2\times 2}(\mathbb{R})$$

Problem 4.2.8 Let $\langle V, +, \cdot, \mathbb{F}\rangle$ be a vector space, and let $U_1 \subseteq V$ and $U_2 \subseteq V$ be vector subspaces. Prove that if $U_1 \cup U_2$ is a vector subspace of V, then either $U_1 \subseteq U_2$ or $U_2 \subseteq U_1$.

Problem 4.2.9 Let V a vector space over a field $\mathbb{F}$. If $\mathbb{F}$ is infinite, show that V is **not** the set-theoretic union of a finite number of **proper** subspaces.

Problem 4.2.10 Give an example of a finite vector space V over a finite field $\mathbb{F}$ such that

$$V = V_1 \cup V_2 \cup V_3,$$

where the V_k are proper subspaces.

4.3 Linear Independence

197 Definition Let $(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \mathbb{F}^n$. Then the vectorial sum

$$\sum_{j=1}^{n} \lambda_j \vec{a}_j$$

is said to be a *linear combination* of the vectors $\vec{a}_i \in V, 1 \leq i \leq n$.

198 Example Any matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{M}_{2\times 2}(\mathbb{R})$ can be written as a linear combination of the matrices

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

for

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

199 Example Any polynomial of degree at most 2, say $a + bx + cx^2 \in \mathbb{R}_2[x]$ can be written as a linear combination of 1, $x - 1$, and $x^2 - x + 2$, for

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

Generalising the notion of two parallel vectors, we have

200 Definition The vectors $\overrightarrow{a}_i \in V, 1 \leq i \leq n$, are *linearly dependent* or *tied* if

$$\exists(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \mathbb{F}^n \setminus \{\mathbf{0}\} \ \textbf{such that} \ \sum_{j=1}^{n} \lambda_j \overrightarrow{a}_j = \overrightarrow{0},$$

that is, if there is a non-trivial linear combination of them adding to the zero vector.

201 Definition The vectors $\overrightarrow{a}_i \in V, 1 \leq i \leq n$, are *linearly independent* or *free* if they are not linearly dependent. That is, if $(\lambda_1, \lambda_2, \cdots, \lambda_n) \in \mathbb{F}^n$ then

$$\sum_{j=1}^{n} \lambda_j \overrightarrow{a}_j = \overrightarrow{0} \implies \lambda_1 = \lambda_2 = \cdots = \lambda_n = 0_{\mathbb{F}}.$$

☞ *A family of vectors is linearly independent if and only if the only linear combination of them giving the zero-vector is the trivial linear combination.*

202 Example

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} \right\}$$

is a tied family of vectors in $\mathbb{R}^3$, since

$$(1) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + (-2) \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} + (1) \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

203 Example Let $\overrightarrow{u}, \overrightarrow{v}$ be linearly independent vectors in some vector space over a field $\mathbb{F}$ with characteristic different from 2. Shew that the two new vectors $\overrightarrow{x} = \overrightarrow{u} - \overrightarrow{v}$ and $\overrightarrow{y} = \overrightarrow{u} + \overrightarrow{v}$ are also linearly independent.

Solution: ► *Assume that* $a(\overrightarrow{u} - \overrightarrow{v}) + b(\overrightarrow{u} + \overrightarrow{v}) = \overrightarrow{0}$. *Then*

$$(a + b)\overrightarrow{u} + (a - b)\overrightarrow{v} = \overrightarrow{0}.$$

Since $\overrightarrow{u}, \overrightarrow{v}$ *are linearly independent, the above coefficients must be 0, that is,* $a + b = 0_{\mathbb{F}}$ *and* $a - b = 0_{\mathbb{F}}$. *But this gives* $2a = 2b = 0_{\mathbb{F}}$, *which implies* $a = b = 0_{\mathbb{F}}$, *if the characteristic of the field is not 2. This proves the linear independence of* $\overrightarrow{u} - \overrightarrow{v}$ *and* $\overrightarrow{u} + \overrightarrow{v}$. ◄

204 Theorem Let $A \in \mathbf{M}_{m \times n}(\mathbb{F})$. Then the columns of A are linearly independent if and only the only solution to the system $AX = \mathbf{0}_m$ is the trivial solution.

Proof: *Let* $A_1, \ldots, A_n$ *be the columns of* A. *Since*

$$x_1 A_1 + x_2 A_2 + \cdots + x_n A_n = AX,$$

the result follows. ❑

205 Theorem Any family

$$\{\overrightarrow{0}, \overrightarrow{u}_1, \overrightarrow{u}_2, \ldots, \overrightarrow{u}_k\}$$

containing the zero-vector is linearly dependent.

Proof: *This follows at once by observing that*

$$1_{\mathbb{F}}\overrightarrow{0} + 0_{\mathbb{F}}\overrightarrow{u}_1 + 0_{\mathbb{F}}\overrightarrow{u}_2 + \ldots, +0_{\mathbb{F}}\overrightarrow{u}_k = \overrightarrow{0}$$

is a non-trivial linear combination of these vectors equalling the zero-vector. ❑

Homework

Problem 4.3.1 Shew that

$$\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$$

forms a free family of vectors in $\mathbb{R}^3$.

Problem 4.3.2 Prove that the set

$$\left\{\begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}, \begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\0\\1\end{bmatrix}\right\}$$

is a linearly independent set of vectors in $\mathbb{R}^4$ and shew that $X = \begin{bmatrix}1\\2\\1\\1\end{bmatrix}$ can be written as a linear combination of these vectors.

Problem 4.3.3 Let $(\overrightarrow{u}, \overrightarrow{v}) \in (\mathbb{R}^n)^2$. Prove that $|\overrightarrow{u}\bullet\overrightarrow{v}| = ||\overrightarrow{u}||||\overrightarrow{v}||$ if and only if $\overrightarrow{u}$ and $\overrightarrow{v}$ are linearly dependent.

Problem 4.3.4 Prove that

$$\left\{\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}, \begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}\right\}$$

is a linearly independent family over $\mathbb{R}$. Write $\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}$ as a linear combination of these matrices.

Problem 4.3.5 Let $\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3, \overrightarrow{v}_4\}$ be a linearly independent family of vectors. Prove that the family

$$\{\overrightarrow{v}_1 + \overrightarrow{v}_2, \overrightarrow{v}_2 + \overrightarrow{v}_3, \overrightarrow{v}_3 + \overrightarrow{v}_4, \overrightarrow{v}_4 + \overrightarrow{v}_1\}$$

is not linearly independent.

Problem 4.3.6 Let $\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3\}$ be linearly independent vectors in $\mathbb{R}^5$. Are the vectors

$$\begin{aligned}\overrightarrow{b}_1 &= 3\overrightarrow{v}_1 + 2\overrightarrow{v}_2 + 4\overrightarrow{v}_3,\\ \overrightarrow{b}_2 &= \overrightarrow{v}_1 + 4\overrightarrow{v}_2 + 2\overrightarrow{v}_3,\\ \overrightarrow{b}_3 &= 9\overrightarrow{v}_1 + 4\overrightarrow{v}_2 + 3\overrightarrow{v}_3,\\ \overrightarrow{b}_4 &= \overrightarrow{v}_1 + 2\overrightarrow{v}_2 + 5\overrightarrow{v}_3,\end{aligned}$$

linearly independent? Prove or disprove!

Problem 4.3.7 Is the family $\{1, \sqrt{2}\}$ linearly independent over $\mathbb{Q}$?

Problem 4.3.8 Is the family $\{1, \sqrt{2}\}$ linearly independent over $\mathbb{R}$?

Problem 4.3.9 Consider the vector space

$$V = \{a + b\sqrt{2} + c\sqrt{3} : (a, b, c) \in \mathbb{Q}^3\}.$$

1. Shew that $\{1, \sqrt{2}, \sqrt{3}\}$ are linearly independent over $\mathbb{Q}$.

2. Express
$$\frac{1}{1-\sqrt{2}} + \frac{2}{\sqrt{12}-2}$$
as a linear combination of $\{1, \sqrt{2}, \sqrt{3}\}$.

Problem 4.3.10 Let f, g, h belong to $C^\infty(\mathbb{R}^{\mathbb{R}})$ (the space of infinitely continuously differentiable real-valued functions defined on the real line) and be given by

$$f(x) = e^x, g(x) = e^{2x}, h(x) = e^{3x}.$$

Shew that f, g, h are linearly independent over $\mathbb{R}$.

Problem 4.3.11 Let f, g, h belong to $C^\infty(\mathbb{R}^{\mathbb{R}})$ be given by

$$f(x) = \cos^2 x, g(x) = \sin^2 x, h(x) = \cos 2x.$$

Shew that f, g, h are linearly dependent over $\mathbb{R}$.

4.4 Spanning Sets

206 Definition A family $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\} \subseteq V$ is said to *span* or *generate* V if every $\vec{v} \in V$ can be written as a linear combination of the $\vec{u}_j$'s.

207 Theorem If $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\} \subseteq V$ spans V, then any superset

$$\{\vec{w}, \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\} \subseteq V$$

also spans V.

Proof: *This follows at once from*

$$\sum_{i=1}^{l} \lambda_i \vec{u}_i = 0_{\mathbb{F}} \vec{w} + \sum_{i=1}^{l} \lambda_i \vec{u}_i.$$

❑

208 Example The family of vectors

$$\left\{ \vec{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \vec{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

spans $\mathbb{R}^3$ since given $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3$ we may write

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = a\vec{i} + b\vec{j} + c\vec{k}.$$

209 Example Prove that the family of vectors

$$\left\{ \vec{t}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{t}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{t}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

spans $\mathbb{R}^3$.

Solution: ▶ *This follows from the identity*

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a-b)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + (b-c)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = (a-b)\vec{t}_1 + (b-c)\vec{t}_2 + c\vec{t}_3.$$

◀

210 Example Since

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = a\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

the set of matrices $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ is a spanning set for $\mathbf{M}_{2\times 2}(\mathbb{R})$.

211 Example The set

$$\{1, x, x^2, x^3, \ldots, x^n, \ldots\}$$

spans $\mathbb{R}[x]$, the set of polynomials with real coefficients and indeterminate x.

212 Definition The *span* of a family of vectors $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\}$ is the set of all finite linear combinations obtained from the u_i's. We denote the span of $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\}$ by

$$\mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right).$$

213 Theorem Let $\langle V, +, \cdot, \mathbb{F}\rangle$ be a vector space. Then

$$\mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right) \subseteq V$$

is a vector subspace of V.

Proof: *Let $\alpha \in \mathbb{F}$ and let*

$$\vec{x} = \sum_{k=1}^{l} a_k \vec{u}_k, \quad \vec{y} = \sum_{k=1}^{l} b_k \vec{u}_k,$$

be in $\mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right)$ *(some of the coefficients might be $0_{\mathbb{F}}$). Then*

$$\vec{x} + \alpha\vec{y} = \sum_{k=1}^{l} (a_k + \alpha b_k)\vec{u}_k \in \mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right),$$

and so $\mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right)$ *is a subspace of V.* ❑

214 Corollary $\mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right) \subseteq V$ is the smallest vector subspace of V (in the sense of set inclusion) containing the set

$$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\}.$$

Proof: *If $W \subseteq V$ is a vector subspace of V containing the set*

$$\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\}$$

then it contains every finite linear combination of them, and hence, it contains $\mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots,\right)$. ❑

215 Example If $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$, $A \in \mathbf{span}\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}\right)$ then A has the form

$$a\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix} + b\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} + c\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} = \begin{bmatrix}a & c\\ c & b\end{bmatrix},$$

i.e., this family spans the set of all symmetric 2×2 matrices over $\mathbb{R}$.

216 Theorem Let V be a vector space over a field $\mathbb{F}$ and let $(\vec{v}, \vec{w}) \in V^2$, $\gamma \in \mathbb{F} \setminus \{0_{\mathbb{F}}\}$. Then

$$\mathbf{span}\left(\vec{v}, \vec{w}\right) = \mathbf{span}\left(\vec{v}, \gamma\vec{w}\right).$$

Proof: *The equality*

$$a\vec{v} + b\vec{w} = a\vec{v} + (b\gamma^{-1})(\gamma\vec{w}),$$

proves the statement. ❑

217 Theorem Let V be a vector space over a field $\mathbb{F}$ and let $(\vec{v}, \vec{w}) \in V^2$, $\gamma \in \mathbb{F}$. Then

$$\mathbf{span}\left(\vec{v}, \vec{w}\right) = \mathbf{span}\left(\vec{w}, \vec{v} + \gamma\vec{w}\right).$$

Proof: *This follows from the equality*

$$a\vec{v} + b\vec{w} = a(\vec{v} + \gamma\vec{w}) + (b - a\gamma)\vec{w}.$$

❑

Homework

Problem 4.4.1 Let $\mathbb{R}_3[x]$ denote the set of polynomials with degree at most 3 and real coefficients. Prove that the set

$$\{1, 1 + x, (1 + x)^2, (1 + x)^3\}$$

spans $\mathbb{R}_3[x]$.

Problem 4.4.2 Shew that $\begin{bmatrix}1\\ 1\\ -1\end{bmatrix} \not\in \mathbf{span}\left(\begin{bmatrix}1\\ 0\\ -1\end{bmatrix}, \begin{bmatrix}0\\ 1\\ -1\end{bmatrix}\right)$.

Problem 4.4.3 What is $\mathbf{span}\left(\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\right)$?

Problem 4.4.4 Prove that

$$\mathbf{span}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}, \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\right) = \mathbf{M}_{2\times 2}(\mathbb{R}).$$

Problem 4.4.5 For the vectors in $\mathbb{R}^3$,

$$\vec{a} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \vec{b} = \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \vec{c} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{d} = \begin{bmatrix} 3 \\ 8 \\ 5 \end{bmatrix},$$

prove that

$$\mathbf{span}\left(\vec{a}, \vec{b}\right) = \mathbf{span}\left(\vec{c}, \vec{d}\right).$$

4.5 Bases

218 Definition A family $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots\} \subseteq V$ is said to be a *basis* of V if (i) they are linearly independent, (ii) they span V.

219 Example The family

$$\vec{e}_i = \begin{bmatrix} 0_{\mathbb{F}} \\ \vdots \\ 0_{\mathbb{F}} \\ 1_{\mathbb{F}} \\ 0_{\mathbb{F}} \\ \vdots \\ 0_{\mathbb{F}} \end{bmatrix},$$

where there is a $1_{\mathbb{F}}$ on the i-th slot and $0_{\mathbb{F}}$'s on the other $n-1$ positions, is a basis for $\mathbb{F}^n$.

220 Theorem Let $\langle V, +, \cdot, \mathbb{F}\rangle$ be a vector space and let

$$U = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k, \ldots\} \subseteq V$$

be a family of linearly independent vectors in V which is maximal in the sense that if U' is any other family of vectors of V properly containing U then U' is a dependent family. Then U forms a basis for V.

Proof: *Since U is a linearly independent family, we need only to prove that it spans V. Take $\vec{v} \in V$. If $\vec{v} \in U$ then there is nothing to prove, so assume that $\vec{v} \in V \setminus U$. Consider the set $U' = U \cup \{\vec{v}\}$. This set properly contains U, and so, by assumption, it forms a dependent family. There exists scalars $\alpha_0, \alpha_1, \ldots, \alpha_n$ such that*

$$\alpha_0\vec{v} + \alpha_1\vec{u}_1 + \cdots + \alpha_n\vec{u}_n = \vec{0}.$$

Now, $\alpha_0 \neq 0_{\mathbb{F}}$, otherwise the $\vec{u}_i$ would be linearly dependent. Hence α_0^{-1} exists and we have

$$\vec{v} = -\alpha_0^{-1}(\alpha_1\vec{u}_1 + \cdots + \alpha_n\vec{u}_n),$$

and so the $\vec{u}_i$ span V. ❑

☞ *From Theorem 220 it follows that to shew that a vector space has a basis it is enough to shew that it has a maximal linearly independent set of vectors. Such a proof requires something called Zörn's Lemma, and it is beyond our scope. We dodge the whole business by taking as an axiom that every vector space possesses a basis.*

221 Theorem (Steinitz Replacement Theorem) Let $\langle V, +, \cdot, \mathbb{F}\rangle$ be a vector space and let $U = \{\vec{u}_1, \vec{u}_2, \ldots\} \subseteq V$. Let $W = \{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$ be an independent family of vectors in **span** (U). Then there exist k of the $\vec{u}_i$'s, say $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ which may be replaced by the $\vec{w}_i$'s in such a way that

$$\mathbf{span}\left(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k, \vec{u}_{k+1}, \ldots\right) = \mathbf{span}\left(U\right).$$

Proof: *We prove this by induction on k. If $k = 1$, then*

$$\vec{w}_1 = \alpha_1\vec{u}_1 + \alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n$$

for some n and scalars α_i. There is an $\alpha_i \neq 0_{\mathbb{F}}$, since otherwise $\vec{w}_1 = \vec{0}$ contrary to the assumption that the $\vec{w}_i$ are linearly independent. After reordering, we may assume that $\alpha_1 \neq 0_{\mathbb{F}}$. Hence

$$\vec{u}_1 = \alpha_1^{-1}(\vec{w}_1 - (\alpha_2\vec{u}_2 + \cdots + \alpha_n\vec{u}_n)),$$

*and so $\vec{u}_1 \in$ **span** $(\vec{w}_1, \vec{u}_2, \ldots,)$ and*

$$\mathbf{span}\left(\vec{w}_1, \vec{u}_2, \ldots,\right) = \mathbf{span}\left(\vec{u}_1, \vec{u}_2, \ldots,\right).$$

Assume now that the theorem is true for any set of fewer than k independent vectors. We may thus assume that that $\{\vec{u}_1, \ldots\}$ has more than $k-1$ vectors and that

$$\mathbf{span}\left(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_{k-1}, \vec{u}_k, , \ldots\right) = \mathbf{span}\left(U\right).$$

Since $\vec{w}_k \in U$ we have

$$\vec{w}_k = \beta_1\vec{w}_1 + \beta_2\vec{w}_2 + \cdots + \beta_{k-1}\vec{w}_{k-1} + \gamma_k\vec{u}_k + \gamma_{k+1}\vec{u}_{k+1} + \gamma_m\vec{u}_m.$$

If all the $\gamma_i = 0_{\mathbb{F}}$, then the $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$ would be linearly dependent, contrary to assumption. Thus there is a $\gamma_i \neq 0_{\mathbb{F}}$, and after reordering, we may assume that $\gamma_k \neq 0_{\mathbb{F}}$. We have therefore

$$\vec{u}_k = \gamma_k^{-1}(\vec{w}_k - (\beta_1\vec{w}_1 + \beta_2\vec{w}_2 + \cdots + \beta_{k-1}\vec{w}_{k-1} + \gamma_{k+1}\vec{u}_{k+1} + \gamma_m\vec{u}_m)).$$

But this means that

$$\mathbf{span}\left(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k, \vec{u}_{k+1}, \ldots\right) = \mathbf{span}\left(U\right).$$

This finishes the proof. ❑

222 Corollary Let $\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n\}$ be an independent family of vectors with $V =$ **span** $(\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n)$. If we also have $V =$ **span** $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_\nu)$, then

1. $n \leq \nu$,
2. $n = \nu$ if and only if the $\{\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_\nu\}$ are a linearly independent family.
3. Any basis for V has exactly n elements.

Proof:

1. *In the Steinitz Replacement Theorem 221 replace the first n $\vec{u}_i$'s by the $\vec{w}_i$'s and $n \leq \nu$ follows.*
2. *If $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_\nu\}$ are a linearly independent family, then we may interchange the rôle of the $\vec{w}_i$ and $\vec{u}_i$ obtaining $\nu \leq n$. Conversely, if $\nu = n$ and if the $\vec{u}_i$ are dependent, we could express some $\vec{u}_i$ as a linear combination of the remaining $\nu-1$ vectors, and thus we would have shewn that some $\nu - 1$ vectors span V. From (1) in this corollary we would conclude that $n \leq \nu - 1$, contradicting $n = \nu$.*
3. *This follows from the definition of what a basis is and from (2) of this corollary.*

❑

223 Definition The *dimension* of a vector space $\langle V, +, \cdot, \mathbb{F}\rangle$ is the number of elements of any of its bases, and we denote it by $\mathbf{dim}\, V$.

224 Theorem Any linearly independent family of vectors

$$\{\overrightarrow{x}_1, \overrightarrow{x}_2, \ldots, \overrightarrow{x}_k\}$$

in a vector space V can be completed into a family

$$\{\overrightarrow{x}_1, \overrightarrow{x}_2, \ldots, \overrightarrow{x}_k, \overrightarrow{y}_{k+1}, \overrightarrow{y}_{k+2}, \ldots\}$$

so that this latter family become a basis for V.

Proof: *Take any basis* $\{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_k, \overrightarrow{u}_{k+1}, \ldots, \}$ *and use Steinitz Replacement Theorem 221.*
❑

225 Corollary If $U \subseteq V$ is a vector subspace of a finite dimensional vector space V then $\mathbf{dim}\, U \leq \mathbf{dim}\, V$.

Proof: *Since any basis of* U *can be extended to a basis of* V*, it follows that the number of elements of the basis of* U *is at most as large as that for* V*.* ❑

226 Example Find a basis and the dimension of the space generated by the set of symmetric matrices in $\mathbf{M}_{n\times n}(\mathbb{R})$.

Solution: ► *Let* $\mathbf{E}_{ij} \in \mathbf{M}_{n\times n}(\mathbb{R})$ *be the* $n \times n$ *matrix with a* 1 *on the* ij*-th position and* 0*'s everywhere else. For* $1 \leq i < j \leq n$*, consider the* $\binom{n}{2} = \frac{n(n-1)}{2}$ *matrices* $A_{ij} = \mathbf{E}_{ij} + \mathbf{E}_{ji}$*. The* A_{ij} *have a* 1 *on the* ij*-th and* ji*-th position and* 0*'s everywhere else. They, together with the* n *matrices* $\mathbf{E}_{ii}, 1 \leq i \leq n$ *constitute a basis for the space of symmetric matrices. The dimension of this space is thus*

$$\frac{n(n-1)}{2} + n = \frac{n(n+1)}{2}.$$

◄

227 Theorem Let $\{\overrightarrow{u}_1, \ldots, \overrightarrow{u}_n\}$ be vectors in $\mathbb{R}^n$. Then the $\overrightarrow{u}$'s form a basis if and only if the $n \times n$ matrix A formed by taking the $\overrightarrow{u}$'s as the columns of A is invertible.

Proof: *Since we have the right number of vectors, it is enough to prove that the* $\overrightarrow{u}$*'s are linearly independent. But if* $X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$*, then*

$$x_1\overrightarrow{u}_1 + \cdots + x_n\overrightarrow{u}_n = AX.$$

If A *is invertible, then* $AX = \mathbf{0}_n \implies X = A^{-1}\mathbf{0}_n = \mathbf{0}_n$*, meaning that* $x_1 = x_2 = \cdots x_n = 0$*, so the* $\overrightarrow{u}$*'s are linearly independent.*

Conversely, assume that the $\overrightarrow{u}$*'s are linearly independent. Then the equation* $AX = \mathbf{0}_n$ *has a unique solution. Let* $r = \mathbf{rank}\,(A)$ *and let* $(P, Q) \in (\mathbf{GL}_n(\mathbb{R}))^2$ *be matrices such that* $A = P^{-1}D_{n,n,r}Q^{-1}$*, where* $D_{n,n,r}$ *is the Hermite normal form of* A*. Thus*

$$AX = \mathbf{0}_n \implies P^{-1}D_{n,n,r}Q^{-1}X = \mathbf{0}_n \implies D_{n,n,r}Q^{-1}X = \mathbf{0}_n.$$

Put $Q^{-1}X = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$. *Then*

$$D_{n,n,r}Q^{-1}X = \mathbf{0}_n \implies y_1\overrightarrow{e}_1 + \cdots + y_r\overrightarrow{e}_r = \mathbf{0}_n,$$

where $\overrightarrow{e}_j$ *is the* n*-dimensional column vector with a* 1 *on the* j*-th slot and* 0*'s everywhere else. If* $r < n$ *then* $y_{r+1}, \ldots, y_n$ *can be taken arbitrarily and so there would not be a unique solution, a contradiction. Hence* $r = n$ *and* A *is invertible.* ❑

Homework

Problem 4.5.1 In problem 4.2.2 we saw that

$$X = \left\{ \begin{bmatrix} a \\ 2a - 3b \\ 5b \\ a + 2b \\ a \end{bmatrix} : a, b \in \mathbb{R} \right\}$$

is a vector subspace of $\mathbb{R}^5$. Find a basis for this subspace.

Problem 4.5.2 Let $\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3, \overrightarrow{v}_4, \overrightarrow{v}_5\}$ be a basis for a vector space V over a field $\mathbb{F}$. Prove that

$$\{\overrightarrow{v}_1 + \overrightarrow{v}_2, \overrightarrow{v}_2 + \overrightarrow{v}_3, \overrightarrow{v}_3 + \overrightarrow{v}_4, \overrightarrow{v}_4 + \overrightarrow{v}_5, \overrightarrow{v}_5 + \overrightarrow{v}_1\}$$

is also a basis for V.

Problem 4.5.3 Find a basis for the solution space of the system of $n + 1$ linear equations of $2n$ unknowns

$$\begin{aligned} x_1 + x_2 + \cdots + x_n &= 0, \\ x_2 + x_3 + \cdots + x_{n+1} &= 0, \\ &\vdots \\ x_{n+1} + x_{n+2} + \cdots + x_{2n} &= 0. \end{aligned}$$

Problem 4.5.4 Prove that the set V of skew-symmetric $n \times n$ matrices is a subspace of $\mathbf{M}_{n\times n}(\mathbb{R})$ and find its dimension. Exhibit a basis for V.

Problem 4.5.5 Prove that the set

$$X = \{(a, b, c, d) | b + 2c = 0\} \subseteq \mathbb{R}^4$$

is a vector subspace of $\mathbb{R}^4$. Find its dimension and a basis for X.

Problem 4.5.6 Prove that the dimension of the vector subspace of lower triangular $n \times n$ matrices is $\dfrac{n(n+1)}{2}$ and find a basis for this space.

Problem 4.5.7 Find a basis and the dimension of

$$X = \textbf{span}\left(\vec{v_1} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix}1\\1\\1\\0\end{bmatrix}, \quad \vec{v_3} = \begin{bmatrix}2\\2\\2\\1\end{bmatrix}\right).$$

Problem 4.5.8 Find a basis and the dimension of

$$X = \textbf{span}\left(\vec{v_1} = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix}1\\1\\1\\0\end{bmatrix}, \quad \vec{v_3} = \begin{bmatrix}2\\2\\2\\2\end{bmatrix}\right).$$

Problem 4.5.9 Find a basis and the dimension of

$$X = \textbf{span}\left(\vec{v_1} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \quad \vec{v_2} = \begin{bmatrix}1 & 0\\2 & 0\end{bmatrix}, \quad \vec{v_3} = \begin{bmatrix}0 & 1\\2 & 0\end{bmatrix}, \vec{v_4} = \begin{bmatrix}1 & -1\\0 & 0\end{bmatrix}\right).$$

Problem 4.5.10 Prove that the set

$$V = \left\{\begin{bmatrix}a & b & c\\0 & d & f\\0 & 0 & g\end{bmatrix} \in \mathbf{M}_{3\times 3}(\mathbb{R}) : a+b+c=0, \quad a+d+g=0\right\}$$

is a vector space of $\mathbf{M}_{3\times 3}(\mathbb{R})$ and find a basis for it and its dimension.

4.6 Coordinates

228 Theorem Let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ be a basis for a vector space V. Then any $\vec{v} \in V$ has a unique representation

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n.$$

Proof: *Let*

$$\vec{v} = b_1\vec{v}_1 + b_2\vec{v}_2 + \cdots + b_n\vec{v}_n$$

be another representation of $\vec{v}$. *Then*

$$\vec{0} = (a_1 - b_1)\vec{v}_1 + (a_2 - b_2)\vec{v}_2 + \cdots + (a_n - b_n)\vec{v}_n.$$

Since $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ *forms a basis for* V, *they are a linearly independent family. Thus we must have*

$$a_1 - b_1 = a_2 - b_2 = \cdots = a_n - b_n = 0_{\mathbb{F}},$$

that is

$$a_1 = b_1; a_2 = b_2; \cdots ; a_n = b_n,$$

proving uniqueness. ❑

229 Definition An *ordered basis* $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of a vector space V is a basis where the order of the $\vec{v}_k$ has been fixed. Given an ordered basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of a vector space V, Theorem 228 ensures that there are unique $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$ such that

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2 + \cdots + a_n\vec{v}_n.$$

The a_k's are called the *coordinates* of the vector $\vec{v}$.

230 Example The standard ordered basis for $\mathbb{R}^3$ is $\mathscr{S} = \{\vec{i}, \vec{j}, \vec{k}\}$. The vector $\begin{bmatrix}1\\2\\3\end{bmatrix} \in \mathbb{R}^3$ for example, has coordinates $(1, 2, 3)_{\mathscr{S}}$. If the order of the basis were changed to the ordered basis $\mathscr{S}_1 = \{\vec{i}, \vec{k}, \vec{j}\}$, then $\begin{bmatrix}1\\2\\3\end{bmatrix} \in \mathbb{R}^3$ would have coordinates $(1, 3, 2)_{\mathscr{S}_1}$.

☞ *Usually, when we give a coordinate representation for a vector $\vec{v} \in \mathbb{R}^n$, we assume that we are using the standard basis.*

231 Example Consider the vector $\begin{bmatrix}1\\2\\3\end{bmatrix} \in \mathbb{R}^3$ (given in standard representation). Since

$$\begin{bmatrix}1\\2\\3\end{bmatrix} = -1\begin{bmatrix}1\\0\\0\end{bmatrix} - 1\begin{bmatrix}1\\1\\0\end{bmatrix} + 3\begin{bmatrix}1\\1\\1\end{bmatrix},$$

under the ordered basis $\mathscr{B}_1 = \left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}1\\1\\0\end{bmatrix}, \begin{bmatrix}1\\1\\1\end{bmatrix}\right\}$, $\begin{bmatrix}1\\2\\3\end{bmatrix}$ has coordinates $(-1, -1, 3)_{\mathscr{B}_1}$. We write

$$\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}-1\\-1\\3\end{bmatrix}_{\mathscr{B}_1}.$$

232 Example The vectors of

$$\mathscr{B}_1 = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}$$

are non-parallel, and so form a basis for $\mathbb{R}^2$. So do the vectors

$$\mathscr{B}_2 = \left\{ \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}.$$

Find the coordinates of $\begin{bmatrix} 3 \\ 4 \end{bmatrix}_{\mathscr{B}_1}$ in the base $\mathscr{B}_2$.

Solution: ▸ *We are seeking* x, y *such that*

$$3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = x \begin{bmatrix} 2 \\ 1 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \end{bmatrix} \implies \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_{\mathscr{B}_2}.$$

Thus

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix}_{\mathscr{B}_2} &= \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & -\frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} \frac{2}{3} & 1 \\ -\frac{1}{3} & -1 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 6 \\ -5 \end{bmatrix}_{\mathscr{B}_2}. \end{aligned}$$

Let us check by expressing both vectors in the standard basis of $\mathbb{R}^2$*:*

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix}_{\mathscr{B}_1} = 3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 4 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 \\ 11 \end{bmatrix},$$

$$\begin{bmatrix} 6 \\ -5 \end{bmatrix}_{\mathscr{B}_2} = 6 \begin{bmatrix} 2 \\ 1 \end{bmatrix} - 5 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 7 \\ 11 \end{bmatrix}.$$

◂

In general let us consider bases $\mathscr{B}_1$, $\mathscr{B}_2$ for the same vector space $\mathbf{V}$. We want to convert $X_{\mathscr{B}_1}$ to $Y_{\mathscr{B}_2}$. We let $\mathbf{A}$ be the matrix formed with the column vectors of $\mathscr{B}_1$ in the given order an $\mathbf{B}$ be the matrix formed with the column vectors of $\mathscr{B}_2$ in the given order. Both $\mathbf{A}$ and $\mathbf{B}$ are invertible matrices since the $\mathscr{B}$'s are bases, in view of Theorem 227. Then we must have

$$\mathbf{A}X_{\mathscr{B}_1} = \mathbf{B}Y_{\mathscr{B}_2} \implies Y_{\mathscr{B}_2} = \mathbf{B}^{-1}\mathbf{A}X_{\mathscr{B}_1}.$$

Also,

$$X_{\mathscr{B}_1} = \mathbf{A}^{-1}\mathbf{B}Y_{\mathscr{B}_2}.$$

This prompts the following definition.

233 Definition Let $\mathscr{B}_1 = \{\overrightarrow{u}_1, \overrightarrow{u}_2, \ldots, \overrightarrow{u}_n\}$ and $\mathscr{B}_2 = \{\overrightarrow{v}_1, \overrightarrow{v}_2, \ldots, \overrightarrow{v}_n\}$ be two ordered bases for a vector space $\mathbf{V}$. Let $\mathbf{A} \in \mathbf{M}_{n\times n}(\mathbb{F})$ be the matrix having the $\overrightarrow{u}$'s as its columns and let $\mathbf{B} \in \mathbf{M}_{n\times n}(\mathbb{F})$ be the matrix having the $\overrightarrow{v}$'s as its columns. The matrix $\mathbf{P} = \mathbf{B}^{-1}\mathbf{A}$ is called the *transition matrix* from $\mathscr{B}_1$ to $\mathscr{B}_2$ and the matrix $\mathbf{P}^{-1} = \mathbf{A}^{-1}\mathbf{B}$ is called the *transition matrix* from $\mathscr{B}_2$ to $\mathscr{B}_1$.

234 Example Consider the bases of $\mathbb{R}^3$

$$\mathscr{B}_1 = \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right\},$$

$$\mathscr{B}_2 = \left\{ \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} \right\}.$$

Find the transition matrix from $\mathscr{B}_1$ to $\mathscr{B}_2$ and also the transition matrix from $\mathscr{B}_2$ to $\mathscr{B}_1$. Also find the coordinates of $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}_{\mathscr{B}_1}$ in terms of $\mathscr{B}_2$.

Solution: ► *Let*

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}.$$

The transition matrix from $\mathscr{B}_1$ to $\mathscr{B}_2$ is

$$\begin{aligned} \mathbf{P} &= \mathbf{B}^{-1}\mathbf{A} \\ &= \begin{bmatrix} 1 & 1 & 2 \\ 1 & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & -1 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & 0 \\ -2 & -1 & -0 \\ 2 & 1 & \frac{1}{2} \end{bmatrix}. \end{aligned}$$

The transition matrix from $\mathscr{B}_2$ to $\mathscr{B}_1$ is

$$\mathbf{P}^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & \frac{1}{2} \end{bmatrix}^{-1} = \begin{bmatrix} -1 & 0 & 0 \\ 2 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}.$$

Now,

$$\mathbf{Y}_{\mathscr{B}_2} = \begin{bmatrix} -1 & 0 & 0 \\ -2 & -1 & 0 \\ 2 & 1 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}_{\mathscr{B}_1} = \begin{bmatrix} -1 \\ -4 \\ \frac{11}{2} \end{bmatrix}_{\mathscr{B}_2}.$$

As a check, observe that in the standard basis for $\mathbb{R}^3$

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}_{\mathscr{B}_1} = 1\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} -1 \\ -4 \\ \frac{11}{2} \end{bmatrix}_{\mathscr{B}_2} = -1\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} - 4\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{11}{2}\begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \\ 1 \end{bmatrix}.$$

◀

Homework

Problem 4.6.1 1. Prove that the following vectors are linearly independent in $\mathbb{R}^4$

$$\vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{a}_4 = \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

2. Find the coordinates of $\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$ under the ordered basis $(\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4)$.

3. Find the coordinates of $\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$ under the ordered basis $(\vec{a}_1, \vec{a}_3, \vec{a}_2, \vec{a}_4)$.

Problem 4.6.2 Consider the matrix

$$\mathbf{A}(a) = \begin{bmatrix} a & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & a & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

❶ Determine all a for which $\mathbf{A}(a)$ is not invertible.

❷ Find the inverse of $\mathbf{A}(a)$ when $\mathbf{A}(a)$ is invertible.

❸ Find the transition matrix from the basis

$$\mathscr{B}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

to the basis

$$\mathscr{B}_2 = \begin{bmatrix} a \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ a \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Chapter 5

Linear Transformations

5.1 Linear Transformations

235 Definition Let $\langle V, +, \cdot, \mathbb{F}\rangle$ and $\langle W, +, \cdot, \mathbb{F}\rangle$ be vector spaces over the same field $\mathbb{F}$. A *linear transformation* or *homomorphism*

$$L : \begin{array}{ccc} V & \to & W \\ \overrightarrow{a} & \mapsto & L(\overrightarrow{a}) \end{array},$$

is a function which is

- **Linear:** $L(\overrightarrow{a} + \overrightarrow{b}) = L(\overrightarrow{a}) + L(\overrightarrow{b})$,
- **Homogeneous:** $L(\alpha\overrightarrow{a}) = \alpha L(\overrightarrow{a})$, for $\alpha \in \mathbb{F}$.

☞ *It is clear that the above two conditions can be summarised conveniently into*

$$L(\overrightarrow{a} + \alpha\overrightarrow{b}) = L(\overrightarrow{a}) + \alpha L(\overrightarrow{b}).$$

236 Example Let

$$L : \begin{array}{ccc} \mathbf{M}_{n\times n}(\mathbb{R}) & \to & \mathbb{R} \\ A & \mapsto & \mathbf{tr}\,(A) \end{array}.$$

Then L is linear, for if $(A, B) \in \mathbf{M}_{n\times n}(\mathbb{R})$, then

$$L(A + \alpha B) = \mathbf{tr}\,(A + \alpha B) = \mathbf{tr}\,(A) + \alpha\mathbf{tr}\,(B) = L(A) + \alpha L(B).$$

237 Example Let

$$L : \begin{array}{ccc} \mathbf{M}_{n\times n}(\mathbb{R}) & \to & \mathbf{M}_{n\times n}(\mathbb{R}) \\ A & \mapsto & A^T \end{array}.$$

Then L is linear, for if $(A, B) \in \mathbf{M}_{n\times n}(\mathbb{R})$, then

$$L(A + \alpha B) = (A + \alpha B)^T = A^T + \alpha B^T = L(A) + \alpha L(B).$$

238 Example For a point $(x, y) \in \mathbb{R}^2$, its reflexion about the y-axis is $(-x, y)$. Prove that

$$R : \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x, y) & \mapsto & (-x, y) \end{array}$$

is linear.

Solution: ► *Let* $(x_1, y_1) \in \mathbb{R}^2$, $(x_2, y_2) \in \mathbb{R}^2$, *and* $\alpha \in \mathbb{R}$. *Then*

$$\begin{aligned} R((x_1, y_1) + \alpha(x_2, y_2)) &= R((x_1 + \alpha x_2, y_1 + \alpha y_2)) \\ &= (-(x_1 + \alpha x_2), y_1 + \alpha y_2) \\ &= (-x_1, y_1) + \alpha(-x_2, y_2) \\ &= R((x_1, y_1)) + \alpha R((x_2, y_2)), \end{aligned}$$

whence R *is linear.* ◄

239 Example Let $L : \mathbb{R}^2 \to \mathbb{R}^4$ be a linear transformation with

$$L \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 3 \end{bmatrix}; \quad L \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ 2 \\ 3 \end{bmatrix}.$$

Find $L \begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

Solution: ► *Since*

$$\begin{bmatrix} 5 \\ 3 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} -1 \\ 1 \end{bmatrix},$$

we have

$$L \begin{bmatrix} 5 \\ 3 \end{bmatrix} = 4L \begin{bmatrix} 1 \\ 1 \end{bmatrix} - L \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} -1 \\ 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -6 \\ 4 \\ 6 \\ 9 \end{bmatrix}.$$

◄

240 Theorem Let $\langle V, +, \cdot, \mathbb{F} \rangle$ and $\langle W, +, \cdot, \mathbb{F} \rangle$ be vector spaces over the same field $\mathbb{F}$, and let $L : V \to W$ be a linear transformation. Then

- $L(\overrightarrow{0}_V) = \overrightarrow{0}_W$.
- $\forall \overrightarrow{x} \in V, L(-\overrightarrow{x}) = -L(\overrightarrow{x})$.

Proof: *We have*
$$L(\overrightarrow{0}_V) = L(\overrightarrow{0}_V + \overrightarrow{0}_V) = L(\overrightarrow{0}_V) + L(\overrightarrow{0}_V),$$
hence
$$L(\overrightarrow{0}_V) - L(\overrightarrow{0}_V) = L(\overrightarrow{0}_V).$$
Since
$$L(\overrightarrow{0}_V) - L(\overrightarrow{0}_V) = \overrightarrow{0}_W,$$
we obtain the first result.

Now
$$\overrightarrow{0}_W = L(\overrightarrow{0}_V) = L(\overrightarrow{x} + (-\overrightarrow{x})) = L(\overrightarrow{x}) + L(-\overrightarrow{x}),$$
from where the second result follows. ❑

Homework

Problem 5.1.1 Consider $L : \mathbb{R}^3 \to \mathbb{R}^3$,
$$L\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y - z \\ x + y + z \\ z \end{bmatrix}.$$
Prove that L is linear.

Problem 5.1.2 Let $A \in \mathbf{GL}_n(\mathbb{R})$ be a fixed matrix. Prove that
$$L : \begin{array}{ccc} \mathbf{M}_{n\times n}(\mathbb{R}) & \to & \mathbf{M}_{n\times n}(\mathbb{R}) \\ H & \mapsto & -A^{-1}HA^{-1} \end{array}$$
is a linear transformation.

Problem 5.1.3 Let V be a vector space and let $S \subseteq V$. The set S is said to be *convex* if $\forall \alpha \in [0;1], \forall \mathbf{x}, \mathbf{y} \in S$, $(1-\alpha)\mathbf{x} + \alpha\mathbf{y} \in S$, that is, for any two points in S, the straight line joining them also belongs to S. Let $T : V \to W$ be a linear transformation from the vector space V to the vector space W. Prove that T maps convex sets into convex sets.

5.2 Kernel and Image of a Linear Transformation

241 Definition Let $\langle V, +, \cdot, \mathbb{F}\rangle$ and $\langle W, +, \cdot, \mathbb{F}\rangle$ be vector spaces over the same field $\mathbb{F}$, and
$$T : \begin{array}{ccc} V & \to & W \\ \overrightarrow{v} & \mapsto & T(\overrightarrow{v}) \end{array}$$
be a linear transformation. The *kernel* of T is the set
$$\mathbf{ker}(T) = \{\overrightarrow{v} \in V : T(\overrightarrow{v}) = \overrightarrow{0}_W\}.$$
The *image* of T is the set
$$\mathbf{Im}(T) = \{\overrightarrow{w} \in w : \exists \overrightarrow{v} \in V \textbf{ such that } T(\overrightarrow{v}) = \overrightarrow{w}\} = T(V).$$

☞ *Since* $T(\overrightarrow{0}_V) = \overrightarrow{0}_W$ *by Theorem 240, we have* $\overrightarrow{0}_V \in \mathbf{ker}(T)$ *and* $\overrightarrow{0}_W \in \mathbf{Im}(T)$.

242 Theorem Let $\langle V, +, \cdot, \mathbb{F}\rangle$ and $\langle W, +, \cdot, \mathbb{F}\rangle$ be vector spaces over the same field $\mathbb{F}$, and

$$T : \begin{array}{ccc} V & \to & W \\ \overrightarrow{v} & \mapsto & T(\overrightarrow{v}) \end{array}$$

be a linear transformation. Then $\mathbf{ker}(T)$ is a vector subspace of V and $\mathbf{Im}(T)$ is a vector subspace of W.

Proof: *Let $(\overrightarrow{v}_1, \overrightarrow{v}_2) \in (\mathbf{ker}(T))^2$ and $\alpha \in \mathbb{F}$. Then $T(\overrightarrow{v}_1) = T(\overrightarrow{v}_2) = \overrightarrow{0}_V$. We must prove that $\overrightarrow{v}_1 + \alpha\overrightarrow{v}_2 \in \mathbf{ker}(T)$, that is, that $T(\overrightarrow{v}_1 + \alpha\overrightarrow{v}_2) = \overrightarrow{0}_W$. But*

$$T(\overrightarrow{v}_1 + \alpha\overrightarrow{v}_2) = T(\overrightarrow{v}_1) + \alpha T(\overrightarrow{v}_2) = \overrightarrow{0}_V + \alpha\overrightarrow{0}_V = \overrightarrow{0}_V$$

proving that $\mathbf{ker}(T)$ is a subspace of V.

Now, let $(\overrightarrow{w}_1, \overrightarrow{w}_2) \in (\mathbf{Im}(T))^2$ and $\alpha \in \mathbb{F}$. Then $\exists(\overrightarrow{v}_1, \overrightarrow{v}_2) \in V^2$ such that $T(\overrightarrow{v}_1) = \overrightarrow{w}_1$ and $T(\overrightarrow{v}_2) = \overrightarrow{w}_2$. We must prove that $\overrightarrow{w}_1 + \alpha\overrightarrow{w}_2 \in \mathbf{Im}(T)$, that is, that $\exists\overrightarrow{v}$ such that $T(\overrightarrow{v}) = \overrightarrow{w}_1 + \alpha\overrightarrow{w}_2$. But

$$\overrightarrow{w}_1 + \alpha\overrightarrow{w}_2 = T(\overrightarrow{v}_1) + \alpha T(\overrightarrow{v}_2) = T(\overrightarrow{v}_1 + \alpha\overrightarrow{v}_2),$$

and so we may take $\overrightarrow{v} = \overrightarrow{v}_1 + \alpha\overrightarrow{v}_2$. This proves that $\mathbf{Im}(T)$ is a subspace of W.

❑

243 Theorem Let $\langle V, +, \cdot, \mathbb{F}\rangle$ and $\langle W, +, \cdot, \mathbb{F}\rangle$ be vector spaces over the same field $\mathbb{F}$, and

$$T : \begin{array}{ccc} V & \to & W \\ \overrightarrow{v} & \mapsto & T(\overrightarrow{v}) \end{array}$$

be a linear transformation. Then T is injective if and only if $\mathbf{ker}(T) = \overrightarrow{0}_V$.

Proof: *Assume that T is injective. Then there is a unique $\overrightarrow{x} \in V$ mapping to $\overrightarrow{0}_W$:*

$$T(\overrightarrow{x}) = \overrightarrow{0}_W.$$

By Theorem 240, $T(\overrightarrow{0}_V) = \overrightarrow{0}_W$, i.e., a linear transformation takes the zero vector of one space to the zero vector of the target space, and so we must have $\overrightarrow{x} = \overrightarrow{0}_V$.

Conversely, assume that $\mathbf{ker}(T) = \{\overrightarrow{0}_V\}$, and that $T(\overrightarrow{x}) = T(\overrightarrow{y})$. We must prove that $\overrightarrow{x} = \overrightarrow{y}$. But

$$\begin{aligned} T(\overrightarrow{x}) = T(\overrightarrow{y}) & \implies T(\overrightarrow{x}) - T(\overrightarrow{y}) = \overrightarrow{0}_W \\ & \implies T(\overrightarrow{x} - \overrightarrow{y}) = \overrightarrow{0}_W \\ & \implies (\overrightarrow{x} - \overrightarrow{y}) \in \mathbf{ker}(T) \\ & \implies \overrightarrow{x} - \overrightarrow{y} = \overrightarrow{0}_V \\ & \implies \overrightarrow{x} = \overrightarrow{y}, \end{aligned}$$

as we wanted to shew. ❑

244 Theorem (Dimension Theorem) Let $\langle V, +, \cdot, \mathbb{F}\rangle$ and $\langle W, +, \cdot, \mathbb{F}\rangle$ be vector spaces of finite dimension over the same field $\mathbb{F}$, and

$$T : \begin{matrix} V & \to & W \\ \overrightarrow{v} & \mapsto & T(\overrightarrow{v}) \end{matrix}$$

be a linear transformation. Then

$$\mathbf{dim}\,\mathbf{ker}\,(T) + \mathbf{dim}\,\mathbf{Im}\,(T) = \mathbf{dim}\,V.$$

Proof: *Let* $\{\overrightarrow{v}_1, \overrightarrow{v}_2, \ldots, \overrightarrow{v}_k\}$ *be a basis for* $\mathbf{ker}\,(T)$. *By virtue of Theorem 224, we may extend this to a basis* $\mathscr{A} = \{\overrightarrow{v}_1, \overrightarrow{v}_2, \ldots, \overrightarrow{v}_k, \overrightarrow{v}_{k+1}, \overrightarrow{v}_{k+2}, \ldots, \overrightarrow{v}_n\}$ *of* V. *Here* $n = \mathbf{dim}\,V$. *We will now shew that* $\mathscr{B} = \{T(\overrightarrow{v}_{k+1}), T(\overrightarrow{v}_{k+2}), \ldots, T(\overrightarrow{v}_n)\}$ *is a basis for* $\mathbf{Im}\,(T)$. *We prove that (i)* $\mathscr{B}$ *spans* $\mathbf{Im}\,(T)$, *and (ii)* $\mathscr{B}$ *is a linearly independent family.*

Let $\overrightarrow{w} \in \mathbf{Im}\,(T)$. *Then* $\exists \overrightarrow{v} \in V$ *such that* $T(\overrightarrow{v}) = \overrightarrow{w}$. *Now since* $\mathscr{A}$ *is a basis for* V *we can write*

$$\overrightarrow{v} = \sum_{i=1}^{n} \alpha_i \overrightarrow{v}_i.$$

Hence

$$\overrightarrow{w} = T(\overrightarrow{v}) = \sum_{i=1}^{n} \alpha_i T(\overrightarrow{v}_i) = \sum_{i=k+1}^{n} \alpha_i T(\overrightarrow{v}_i),$$

since $T(\overrightarrow{v}_i) = \overrightarrow{0}_V$ *for* $1 \leq i \leq k$. *Thus* $\mathscr{B}$ *spans* $\mathbf{Im}\,(T)$.

To prove the linear independence of the $\mathscr{B}$ *assume that*

$$\overrightarrow{0}_W = \sum_{i=k+1}^{n} \beta_i T(\overrightarrow{v}_i) = T\left(\sum_{i=k+1}^{n} \beta_i \overrightarrow{v}_i\right).$$

This means that $\sum_{i=k+1}^{n} \beta_i \overrightarrow{v}_i \in \mathbf{ker}\,(T)$, *which is impossible unless* $\beta_{k+1} = \beta_{k+2} = \cdots = \beta_n = 0_{\mathbb{F}}$.
❑

245 Corollary If $\mathbf{dim}\,V = \mathbf{dim}\,W < +\infty$, then T is injective if and only if it is surjective.

Proof: *Simply observe that if* T *is injective then* $\mathbf{dim}\,\mathbf{ker}\,(T) = 0$, *and if* T *is surjective* $\mathbf{Im}\,(T) = T(V) = W$ *and* $\mathbf{Im}\,(T) = \mathbf{dim}\,W$. ❑

246 Example Let

$$L : \begin{matrix} \mathbf{M}_{2\times 2}(\mathbb{R}) & \to & \mathbf{M}_{2\times 2}(\mathbb{R}) \\ A & \mapsto & A^T - A \end{matrix}.$$

Observe that L is linear. Determine $\mathbf{ker}\,(L)$ and $\mathbf{Im}\,(L)$.

Solution: ▶ *Put* $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *and assume* $L(A) = \mathbf{0}_2$. *Then*

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = L(A) = \begin{bmatrix} a & c \\ b & d \end{bmatrix} - \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (c - b)\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

This means that $c = b$. *Thus*

$$\mathbf{ker}\,(L) = \left\{ \begin{bmatrix} a & b \\ b & d \end{bmatrix} : (a, b, d) \in \mathbb{R}^3 \right\},$$

$$\mathbf{Im}\,(L) = \left\{ \begin{bmatrix} 0 & k \\ -k & 0 \end{bmatrix} : k \in \mathbb{R} \right\}.$$

This means that $\mathbf{dim}\,\mathbf{ker}\,(L) = 3$, *and so* $\mathbf{dim}\,\mathbf{Im}\,(L) = 4 - 3 = 1$. ◀

247 Example Consider the linear transformation $L : \mathbf{M}_{2\times 2}(\mathbb{R}) \to \mathbb{R}_3[X]$ given by

$$L \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a + b)X^2 + (a - b)X^3.$$

Determine $\mathbf{ker}\,(L)$ and $\mathbf{Im}\,(L)$.

Solution: ▶ *We have*

$$0 = L \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a + b)X^2 + (a - b)X^3 \implies a + b = 0,\ a - b = 0, \implies a = b = 0.$$

Thus

$$\mathbf{ker}\,(L) = \left\{ \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} : (c, d) \in \mathbb{R}^2 \right\}.$$

Thus $\mathbf{dim}\,\mathbf{ker}\,(L) = 2$ *and hence* $\mathbf{dim}\,\mathbf{Im}\,(L) = 2$. *Now*

$$(a + b)X^2 + (a - b)X^3 \implies a(X^2 + X^3) + b(X^2 - X^3).$$

Clearly $X^2 + X^3$, *and* $X^2 - X^3$ *are linearly independent and span* $\mathbf{Im}\,(L)$. *Thus*

$$\mathbf{Im}\,(L) = \mathbf{span}\left(X^2 + X^3, X^2 - X^3\right).$$

◀

Homework

Problem 5.2.1 In problem 5.1.1 we saw that $L : \mathbb{R}^3 \to \mathbb{R}^3$,

$$L \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y - z \\ x + y + z \\ z \end{bmatrix}$$

is linear. Determine $\mathbf{ker}\,(L)$ and $\mathbf{Im}\,(L)$.

Problem 5.2.2 Consider the function $L : \mathbb{R}^4 \to \mathbb{R}^2$ given by

$$L \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x + y \\ x - y \end{bmatrix}.$$

1. Prove that L is linear.
2. Determine $\mathbf{ker}(L)$.
3. Determine $\mathbf{Im}(L)$.

Problem 5.2.3 Let

$$L : \begin{array}{ccc} \mathbb{R}^3 & \to & \mathbb{R}^4 \\ \overrightarrow{a} & \mapsto & L(\overrightarrow{a}) \end{array}$$

satisfy

$$L\begin{bmatrix}1\\0\\0\end{bmatrix} = \begin{bmatrix}1\\0\\-1\\0\end{bmatrix}; \quad L\begin{bmatrix}1\\1\\0\end{bmatrix} = \begin{bmatrix}2\\-1\\0\\0\end{bmatrix}; \quad L\begin{bmatrix}0\\0\\1\end{bmatrix} = \begin{bmatrix}1\\-1\\1\\0\end{bmatrix}.$$

Determine $\mathbf{ker}(L)$ and $\mathbf{Im}(L)$.

Problem 5.2.4 It is easy to see that $L : \mathbb{R}^2 \to \mathbb{R}^3$,

$$L\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x+2y\\x+2y\\0\end{bmatrix}$$

is linear. Determine $\mathbf{ker}(L)$ and $\mathbf{Im}(L)$.

Problem 5.2.5 It is easy to see that $L : \mathbb{R}^2 \to \mathbb{R}^3$,

$$L\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x-y\\x+y\\0\end{bmatrix}$$

is linear. Determine $\mathbf{ker}(T)$ and $\mathbf{Im}(T)$.

Problem 5.2.6 It is easy to see that $L : \mathbb{R}^3 \to \mathbb{R}^2$,

$$L\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}x-y-z\\y-2z\end{bmatrix}$$

is linear. Determine $\mathbf{ker}(L)$ and $\mathbf{Im}(L)$.

Problem 5.2.7 Let

$$L : \begin{array}{ccc} \mathbf{M}_{2\times 2}(\mathbb{R}) & \to & \mathbb{R} \\ A & \mapsto & \mathbf{tr}(A) \end{array}.$$

Determine $\mathbf{ker}(L)$ and $\mathbf{Im}(L)$.

Problem 5.2.8

1. Demonstrate that

$$L : \begin{array}{ccc} \mathbf{M}_{2\times 2}(\mathbb{R}) & \to & \mathbf{M}_{2\times 2}(\mathbb{R}) \\ A & \mapsto & A^T + A \end{array}$$

is a linear transformation.

2. Find a basis for $\mathbf{ker}(L)$ and find $\mathbf{dim}\,\mathbf{ker}(L)$
3. Find a basis for $\mathbf{Im}(L)$ and find $\mathbf{dim}\,\mathbf{Im}(L)$.

Problem 5.2.9 Let V be an n-dimensional vector space, where the characteristic of the underlying field is different from 2. A linear transformation $T : V \to V$ is *idempotent* if $T^2 = T$. Prove that if T is idempotent, then

❶ $I - T$ is idempotent (I is the identity function).

❷ $I + T$ is invertible.

❸ $\mathbf{ker}(T) = \mathbf{Im}(I - T)$

5.3 Matrix Representation

Let V, W be two vector spaces over the same field $\mathbb{F}$. Assume that $\mathbf{dim}\, V = m$ and $\{\overrightarrow{v}_i\}_{i\in[1;m]}$ is an ordered basis for V, and that $\mathbf{dim}\, W = n$ and $\mathscr{A} = \{\overrightarrow{w}_i\}_{i\in[1;n]}$ an ordered basis for W. Then

$$\begin{array}{lllll}
L(\overrightarrow{v}_1) & = & a_{11}\overrightarrow{w}_1 + a_{21}\overrightarrow{w}_2 + \cdots + a_{n1}\overrightarrow{w}_n & = & \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix}_{\mathscr{A}} \\
L(\overrightarrow{v}_2) & = & a_{12}\overrightarrow{w}_1 + a_{22}\overrightarrow{w}_2 + \cdots + a_{n2}\overrightarrow{w}_n & = & \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{bmatrix}_{\mathscr{A}} \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
L(\overrightarrow{v}_m) & = & a_{1m}\overrightarrow{w}_1 + a_{2m}\overrightarrow{w}_2 + \cdots + a_{nm}\overrightarrow{w}_n & = & \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{nm} \end{bmatrix}_{\mathscr{A}}
\end{array}.$$

248 Definition The $n \times m$ matrix

$$M_L = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{12} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

formed by the column vectors above is called the *matrix representation of the linear map* L *with respect to the bases* $\{\overrightarrow{v}_i\}_{i\in[1;m]}, \{\overrightarrow{w}_i\}_{i\in[1;n]}$.

249 Example Consider $L : \mathbb{R}^3 \to \mathbb{R}^3$,

$$L\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y - z \\ x + y + z \\ z \end{bmatrix}.$$

Clearly L is a linear transformation.

1. Find the matrix corresponding to L under the standard ordered basis.

2. Find the matrix corresponding to $\mathbf{L}$ under the ordered basis $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, for both the domain and the image of $\mathbf{L}$.

Solution: ▶

1. *The matrix will be a* 3×3 *matrix. We have* $\mathbf{L}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{L}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$, *and* $\mathbf{L}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$, *whence the desired matrix is*
$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. *Call this basis* $\mathscr{A}$. *We have*
$$\mathbf{L}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}_{\mathscr{A}},$$
$$\mathbf{L}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} = -2\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 0\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \\ 0 \end{bmatrix}_{\mathscr{A}},$$
and
$$\mathbf{L}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} = -3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}_{\mathscr{A}},$$
whence the desired matrix is
$$\begin{bmatrix} 0 & -2 & -3 \\ 1 & 2 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

◀

250 Example Let $\mathbb{R}_n[x]$ denote the set of polynomials with real coefficients with degree at most n.

1. Prove that

$$L : \begin{array}{ccc} \mathbb{R}_3[x] & \to & \mathbb{R}_1[x] \\ p(x) & \mapsto & p''(x) \end{array}$$

is a linear transformation. Here $p''(x)$ denotes the second derivative of $p(x)$ with respect to x.

2. Find the matrix of L using the ordered bases $\{1, x, x^2, x^3\}$ for $\mathbb{R}_3[x]$ and $\{1, x\}$ for $\mathbb{R}_2[x]$.

3. Find the matrix of L using the ordered bases $\{1, x, x^2, x^3\}$ for $\mathbb{R}_3[x]$ and $\{1, x+2\}$ for $\mathbb{R}_1[x]$.

4. Find a basis for **ker** (L) and find **dim ker** (L).

5. Find a basis for **Im** (L) and find **dim Im** (L).

Solution: ▶

1. *Let* $(p(x), q(x)) \in (\mathbb{R}_3[x])^2$ *and* $\alpha \in \mathbb{R}$. *Then*

$$L(p(x) + \alpha q(x)) = (p(x) + \alpha q(x))'' = p''(x) + \alpha q''(x) = L(p(x)) + \alpha L(q(x)),$$

whence L *is linear.*

2. *We have*

$$\begin{array}{lllllllll} L(1) & = & \frac{\mathbf{d}^2}{\mathbf{d}x^2}1 & = & 0 & = & 0(1) + 0(x) & = & \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ L(x) & = & \frac{\mathbf{d}^2}{\mathbf{d}x^2}x & = & 0 & = & 0(1) + 0(x) & = & \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ L(x^2) & = & \frac{\mathbf{d}^2}{\mathbf{d}x^2}x^2 & = & 2 & = & 2(1) + 0(x) & = & \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \\ L(x^3) & = & \frac{\mathbf{d}^2}{\mathbf{d}x^2}x^3 & = & 6x & = & 0(1) + 6(x) & = & \begin{bmatrix} 0 \\ 6 \end{bmatrix}, \end{array}$$

whence the matrix representation of L *under the standard basis is*

$$\begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}.$$

3. *We have*

$$\begin{array}{rcccccccl} L(1) & = & \frac{d^2}{dx^2}1 & = & 0 & = & 0(1)+0(x+2) & = & \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ L(x) & = & \frac{d^2}{dx^2}x & = & 0 & = & 0(1)+0(x+2) & = & \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ L(x^2) & = & \frac{d^2}{dx^2}x^2 & = & 2 & = & 2(1)+0(x+2) & = & \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \\ L(x^3) & = & \frac{d^2}{dx^2}x^3 & = & 6x & = & -12(1)+6(x+2) & = & \begin{bmatrix} -12 \\ 6 \end{bmatrix}, \end{array}$$

whence the matrix representation of L *under the standard basis is*

$$\begin{bmatrix} 0 & 0 & 2 & -12 \\ 0 & 0 & 0 & 6 \end{bmatrix}.$$

4. *Assume that* $p(x) = a + bx + cx^2 + dx^3 \in \mathbf{ker}\,(L)$. *Then*

$$0 = L(p(x)) = 2c + 6dx, \quad \forall x \in \mathbb{R}.$$

This means that $c = d = 0$. *Thus* a, b *are free and*

$$\mathbf{ker}\,(L) = \{a + bx : (a, b) \in \mathbb{R}^2\}.$$

Hence $\mathbf{dim\,ker}\,(L) = 2$.

5. *By the Dimension Theorem,* $\mathbf{dim\,Im}\,(L) = 4 - 2 = 2$. *Put* $q(x) = \alpha + \beta x + \gamma x^2 + \delta x^3$. *Then*

$$L(q(x)) = 2\gamma + 6\delta(x) = (2\gamma)(1) + (6\delta)(x).$$

Clearly $\{1, x\}$ *are linearly independent and span* $\mathbf{Im}\,(L)$. *Hence*

$$\mathbf{Im}\,(L) = \mathbf{span}\,(1, x) = \mathbb{R}_1[x].$$

◀

251 Example

1. A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ is such that

$$T(\vec{i}) = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}; \quad T(\vec{j}) = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix}.$$

It is known that

$$\mathbf{Im}\,(T) = \mathbf{span}\left(T(\vec{i}), T(\vec{j})\right)$$

and that

$$\mathbf{ker}(T) = \mathbf{span}\left(\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}\right).$$

Argue that there must be λ and μ such that

$$T(\vec{k}) = \lambda T(\vec{i}) + \mu T(\vec{j}).$$

2. Find λ and μ, and hence, the matrix representing T under the standard ordered basis.

Solution: ▶

1. Since $T(\vec{k}) \in \mathbf{Im}(T)$ and $\mathbf{Im}(T)$ is generated by $T(\vec{i})$ and $T(\vec{k})$ there must be $(\lambda, \mu) \in \mathbb{R}^2$ with

$$T(\vec{k}) = \lambda T(\vec{i}) + \mu T(\vec{j}) = \lambda \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + \mu \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 2\lambda + 3\mu \\ \lambda \\ \lambda - \mu \end{bmatrix}.$$

2. The matrix of T is

$$\begin{bmatrix} T(\vec{i}) & T(\vec{j}) & T(\vec{k}) \end{bmatrix} = \begin{bmatrix} 2 & 3 & 2\lambda + 3\mu \\ 1 & 0 & \lambda \\ 1 & -1 & \lambda - \mu \end{bmatrix}.$$

Since $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \in \mathbf{ker}(T)$ *we must have*

$$\begin{bmatrix} 2 & 3 & 2\lambda + 3\mu \\ 1 & 0 & \lambda \\ 1 & -1 & \lambda - \mu \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Solving the resulting system of linear equations we obtain $\lambda = 1, \mu = 2$. The required matrix is thus

$$\begin{bmatrix} 2 & 3 & 8 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}.$$

◀

☞ *If the linear mapping $L : V \to W$, $\mathbf{dim}\, V = n$, $\mathbf{dim}\, W = m$ has matrix representation $A \in \mathbf{M}_{m\times n}(\mathbb{F})$, then $\mathbf{dim}\,\mathbf{Im}(L) = \mathbf{rank}(A)$.*

Homework

Problem 5.3.1 Let $T : \mathbb{R}^4 \to \mathbb{R}^3$ be a linear transformations such that

$$T\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\1\end{bmatrix}, \quad T\begin{bmatrix}1\\0\\1\\0\end{bmatrix} = \begin{bmatrix}1\\1\\-1\end{bmatrix}, \quad T\begin{bmatrix}1\\1\\1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\-1\end{bmatrix}, \quad T\begin{bmatrix}-1\\-2\\0\\0\end{bmatrix} = \begin{bmatrix}1\\1\\1\end{bmatrix}.$$

Find the matrix of T with respect to the canonical bases. Find the dimensions and describe $\mathbf{ker}\,(T)$ and $\mathbf{Im}\,(T)$.

Problem 5.3.2
1. A linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ has as image the plane with equation $x + y + z = 0$ and as kernel the line $x = y = z$. If

$$T\begin{bmatrix}1\\1\\2\end{bmatrix} = \begin{bmatrix}a\\0\\1\end{bmatrix}, \quad T\begin{bmatrix}2\\1\\1\end{bmatrix} = \begin{bmatrix}3\\b\\-5\end{bmatrix}, \quad T\begin{bmatrix}1\\2\\1\end{bmatrix} = \begin{bmatrix}-1\\2\\c\end{bmatrix}.$$

   Find a, b, c.
2. Find the matrix representation of T under the standard basis.

Problem 5.3.3
1. Prove that $T : \mathbb{R}^2 \to \mathbb{R}^3$

$$T\begin{bmatrix}x\\y\end{bmatrix} = \begin{bmatrix}x+y\\x-y\\2x+3y\end{bmatrix}$$

   is a linear transformation.
2. Find a basis for $\mathbf{ker}\,(T)$ and find $\mathbf{dim}\,\mathbf{ker}\,(T)$
3. Find a basis for $\mathbf{Im}\,(T)$ and find $\mathbf{dim}\,\mathbf{Im}\,(T)$.
4. Find the matrix of T under the ordered bases $\mathscr{A} = \left\{\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}1\\3\end{bmatrix}\right\}$ of $\mathbb{R}^2$ and $\mathscr{B} = \left\{\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\-1\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$ of $\mathbb{R}^3$.

Problem 5.3.4 Let

$$L : \begin{array}{ccc}\mathbb{R}^3 & \to & \mathbb{R}^2\\ \overrightarrow{a} & \mapsto & L(\overrightarrow{a})\end{array},$$

where

$$L\begin{bmatrix}x\\y\\z\end{bmatrix} = \begin{bmatrix}x+2y\\3x-z\end{bmatrix}.$$

Clearly L is linear. Find a matrix representation for L if

1. The bases for both $\mathbb{R}^3$ and $\mathbb{R}^2$ are both the standard ordered bases.

2. The ordered basis for $\mathbb{R}^3$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbb{R}^2$ has the standard ordered basis .

3. The ordered basis for $\mathbb{R}^3$ is $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and the ordered basis for $\mathbb{R}^2$ is $\mathscr{A} = \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$.

Problem 5.3.5 A linear transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ satisfies $\mathbf{ker}(T) = \mathbf{Im}(T)$, and $T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Find the matrix representing T under the standard ordered basis.

Problem 5.3.6 Find the matrix representation for the linear map

$$L : \begin{array}{ccc} \mathbf{M}_{2\times 2}(\mathbb{R}) & \to & \mathbb{R} \\ A & \mapsto & \mathbf{tr}(A) \end{array},$$

under the standard basis

$$\mathscr{A} = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$$

for $\mathbf{M}_{2\times 2}(\mathbb{R})$.

Problem 5.3.7 Let $A \in \mathbf{M}_{n\times p}(\mathbb{R})$, $B \in \mathbf{M}_{p\times q}(\mathbb{R})$, and $C \in \mathbf{M}_{q\times r}(\mathbb{R})$, be such that $\mathbf{rank}(B) = \mathbf{rank}(AB)$. Shew that

$$\mathbf{rank}(BC) = \mathbf{rank}(ABC).$$

Chapter 6

Determinants

6.1 Permutations

252 Definition Let S be a finite set with $n \geq 1$ elements. A *permutation* is a bijective function $\tau : S \to S$. It is easy to see that there are $n!$ permutations from S onto itself.

Since we are mostly concerned with the *action* that τ exerts on S rather than with the particular names of the elements of S, we will take S to be the set $S = \{1, 2, 3, \ldots, n\}$. We indicate a permutation τ by means of the following convenient diagram

$$\tau = \begin{bmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{bmatrix}.$$

253 Definition The notation S_n will denote the set of all permutations on $\{1, 2, 3, \ldots, n\}$. Under this notation, the composition of two permutations $(\tau, \sigma) \in S_n^2$ is

$$\begin{aligned} \tau \circ \sigma &= \begin{bmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{bmatrix} \circ \begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & \cdots & n \\ (\tau \circ \sigma)(1) & (\tau \circ \sigma)(2) & \cdots & (\tau \circ \sigma)(n) \end{bmatrix} \end{aligned}.$$

The k-fold composition of τ is

$$\underbrace{\tau \circ \cdots \circ \tau}_{k \text{ compositions}} = \tau^k.$$

☞ *We usually do away with the $\circ$ and write $\tau \circ \sigma$ simply as $\tau\sigma$. This "product of permutations" is thus simply function composition.*

Given a permutation $\tau : S \to S$, since τ is bijective,

$$\tau^{-1} : S \to S$$

exists and is also a permutation. In fact if

$$\tau = \begin{bmatrix} 1 & 2 & \cdots & n \\ \tau(1) & \tau(2) & \cdots & \tau(n) \end{bmatrix},$$

then

$$\tau^{-1} = \begin{bmatrix} \tau(1) & \tau(2) & \cdots & \tau(n) \\ 1 & 2 & \cdots & n \end{bmatrix}.$$

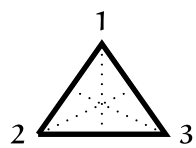


Figure 6.1: S_3 are rotations and reflexions.

254 Example The set S_3 has $3! = 6$ elements, which are given below.

1. **Id** : $\{1, 2, 3\} \to \{1, 2, 3\}$ with
$$\mathbf{Id} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix}.$$

2. $\tau_1 : \{1, 2, 3\} \to \{1, 2, 3\}$ with
$$\tau_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}.$$

3. $\tau_2 : \{1, 2, 3\} \to \{1, 2, 3\}$ with
$$\tau_2 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix}.$$

4. $\tau_3 : \{1, 2, 3\} \to \{1, 2, 3\}$ with
$$\tau_3 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix}.$$

5. $\sigma_1 : \{1, 2, 3\} \to \{1, 2, 3\}$ with
$$\sigma_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}.$$

6. $\sigma_2 : \{1,2,3\} \to \{1,2,3\}$ with

$$\sigma_2 = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}.$$

255 Example The compositions $\tau_1 \circ \sigma_1$ and $\sigma_1 \circ \tau_1$ can be found as follows.

$$\tau_1 \circ \sigma_1 = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} \circ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} = \tau_2.$$

(We read from right to left $1 \to 2 \to 3$ ("1 goes to 2, 2 goes to 3, so 1 goes to 3"), etc. Similarly

$$\sigma_1 \circ \tau_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} \circ \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} = \tau_3.$$

Observe in particular that $\sigma_1 \circ \tau_1 \neq \tau_1 \circ \sigma_1$. Finding all the other products we deduce the following "multiplication table" (where the "multiplication" operation is really composition of functions).

$\circ$	**Id**	τ_1	τ_2	τ_3	σ_1	σ_2
Id	**Id**	τ_1	τ_2	τ_3	σ_1	σ_2
τ_1	τ_1	**Id**	σ_1	σ_2	τ_2	τ_3
τ_2	τ_2	σ_2	**Id**	σ_1	τ_3	τ_1
τ_3	τ_3	σ_1	σ_2	**Id**	τ_1	τ_2
σ_2	σ_2	τ_2	τ_3	τ_1	**Id**	σ_1
σ_1	σ_1	τ_3	τ_1	τ_2	σ_2	**Id**

The permutations in example 254 can be conveniently interpreted as follows. Consider an equilateral triangle with vertices labelled 1, 2 and 3, as in figure 6.1. Each τ_a is a reflexion ("flipping") about the line joining the vertex a with the midpoint of the side opposite a. For example τ_1 fixes 1 and flips 2 and 3. Observe that two successive flips return the vertices to their original position and so $(\forall a \in \{1,2,3\})(\tau_a^2 = \mathbf{Id})$. Similarly, σ_1 is a rotation of the vertices by an angle of $120°$. Three successive rotations restore the vertices to their original position and so $\sigma_1^3 = \mathbf{Id}$.

256 Example To find τ_1^{-1} take the representation of τ_1 and exchange the rows:

$$\tau_1^{-1} = \begin{bmatrix} 1 & 3 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

This is more naturally written as

$$\tau_1^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}.$$

Observe that $\tau_1^{-1} = \tau_1$.

257 Example To find σ_1^{-1} take the representation of σ_1 and exchange the rows:

$$\sigma_1^{-1} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix}.$$

This is more naturally written as

$$\sigma_1^{-1} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix}.$$

Observe that $\sigma_1^{-1} = \sigma_2$.

6.2 Cycle Notation

We now present a shorthand notation for permutations by introducing the idea of a *cycle*. Consider in S_9 the permutation

$$\tau = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 1 & 3 & 6 & 9 & 7 & 8 & 4 & 5 \end{bmatrix}.$$

We start with 1. Since 1 goes to 2 and 2 goes back to 1, we write (12). Now we continue with 3. Since 3 goes to 3, we write (3). We continue with 4. As 4 goes 6, 6 goes to 7, 7 goes 8, and 8 goes back to 4, we write (4678). We consider now 5 which goes to 9 and 9 goes back to 5, so we write (59). We have written τ as a product of disjoint cycles

$$\tau = (12)(3)(4678)(59).$$

This prompts the following definition.

258 Definition Let $l \geq 1$ and let $i_1, \ldots, i_l \in \{1, 2, \ldots n\}$ be distinct. We write $(i_1\ i_2\ \ldots\ i_l)$ for the element $\sigma \in S_n$ such that $\sigma(i_r) = i_{r+1},\ \ 1 \leq r < l$, $\sigma(i_l) = i_1$ and $\sigma(i) = i$ for $i \not\in \{i_1, \ldots, i_l\}$. We say that $(i_1\ i_2\ \ldots\ i_l)$ is a *cycle of length* l. The *order* of a cycle is its length. Observe that if τ has order l then $\tau^l = \mathbf{Id}$.

Observe that $(i_2\ \ldots\ i_l\ i_1) = (i_1\ \ldots\ i_l)$ *etc., and that* $(1) = (2) = \cdots = (n) = \mathbf{Id}$ *. In fact, we have*

$$(i_1\ \ldots\ i_l) = (j_1\ \ldots\ j_m)$$

if and only if (1) $l = m$ *and if (2)* $l > 1$*:* $\exists a$ *such that* $\forall k$*:* $i_k = j_{k+a \mod l}$*. Two cycles* $(i_1, \ldots, i_l)$ *and* $(j_1, \ldots, j_m)$ *are disjoint if* $\{i_1, \ldots, i_l\} \cap \{j_1, \ldots, j_m\} = \varnothing$*. Disjoint cycles commute and if* $\tau = \sigma_1\sigma_2\cdots\sigma_t$ *is the product of disjoint cycles of length* $l_1, l_2, \ldots, l_t$ *respectively, then* τ *has order*

$$\mathbf{lcm}\,(l_1, l_2, \ldots, l_t)\,.$$

259 Example A cycle decomposition for $\alpha \in S_9$,

$$\alpha = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 8 & 7 & 6 & 2 & 3 & 4 & 5 & 9 \end{bmatrix}$$

is

$$(285)(3746).$$

The order of α is $\mathbf{lcm}\,(3, 4) = 12$.

260 Example The cycle decomposition $\beta = (123)(567)$ in S_9 arises from the permutation

$$\beta = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 1 & 4 & 6 & 7 & 5 & 8 & 9 \end{bmatrix}.$$

Its order is $\mathbf{lcm}\,(3,3) = 3$.

261 Example Find a shuffle of a deck of 13 cards that requires 42 repeats to return the cards to their original order.

> **Solution:** ▶ *Here is one (of many possible ones). Observe that* $7+6=13$ *and* $7\times 6 = 42$. *We take the permutation*
>
> $$(1\ 2\ 3\ 4\ 5\ 6\ 7)(8\ 9\ 10\ 11\ 12\ 13)$$
>
> *which has order 42. This corresponds to the following shuffle: For*
>
> $$i \in \{1,2,3,4,5,6,8,9,10,11,12\},$$
>
> *take the* i*th card to the* $(i+1)$*th place, take the* 7*th card to the first position and the 13th card to the* 8*th position. Query: Of all possible shuffles of 13 cards, which one takes the longest to restitute the cards to their original position?* ◀

262 Example Let a shuffle of a deck of 10 cards be made as follows: The top card is put at the bottom, the deck is cut in half, the bottom half is placed on top of the top half, and then the resulting bottom card is put on top. How many times must this shuffle be repeated to get the cards in the initial order? Explain.

> **Solution:** ▶ *Putting the top card at the bottom corresponds to*
>
> $$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 1 \end{bmatrix}.$$
>
> *Cutting this new arrangement in half and putting the lower half on top corresponds to*
>
> $$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 8 & 9 & 10 & 1 & 2 & 3 & 4 & 5 & 6 \end{bmatrix}.$$
>
> *Putting the bottom card of this new arrangement on top corresponds to*
>
> $$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 6 & 7 & 8 & 9 & 10 & 1 & 2 & 3 & 4 & 5 \end{bmatrix} = (1\ 6)(2\ 7)(3\ 8)(4\ 9)(5\ 10).$$
>
> *The order of this permutation is* $\mathbf{lcm}(2,2,2,2,2) = 2$, *so in 2 shuffles the cards are restored to their original position.* ◀

The above examples illustrate the general case, given in the following theorem.

263 Theorem Every permutation in S_n can be written as a product of disjoint cycles.

Proof: *Let* $\tau \in S_n$, $a_1 \in \{1, 2, \ldots, n\}$. *Put* $\tau^k(a_1) = a_{k+1}, k \geq 0$. *Let* $a_1, a_2, \ldots, a_s$ *be the longest chain with no repeats. Then we have* $\tau(a_s) = a_1$. *If the* $\{a_1, a_2, \ldots, a_s\}$ *exhaust* $\{1, 2, \ldots, n\}$ *then we have* $\tau = (a_1\ a_2\ \ldots\ a_s)$. *If not, there exist* $b_1 \in \{1, 2, \ldots, n\} \setminus \{a_1, a_2, \ldots, a_s\}$. *Again, we find the longest chain of distinct* $b_1, b_2, \ldots, b_t$ *such that* $\tau(b_k) = b_{k+1}, k = 1, \ldots, t-1$ *and* $\tau(b_t) = b_1$. *If the* $\{a_1, a_2, \ldots, a_s, b_1, b_2, \ldots, b_t\}$ *exhaust all the* $\{1, 2, \ldots, n\}$ *we have* $\tau = (a_1\ a_2\ \ldots\ a_s)(b_1\ b_2\ \ldots\ b_t)$. *If not we continue the process and find*

$$\tau = (a_1\ a_2\ \ldots\ a_s)(b_1\ b_2\ \ldots\ b_t)(c_1 \ldots) \ldots.$$

This process stops because we have only n *elements.* ❑

264 Definition A *transposition* is a cycle of length 2.[1]

265 Example The cycle (23468) can be written as a product of transpositions as follows

$$(23468) = (28)(26)(24)(23).$$

Notice that this decomposition as the product of transpositions is not unique. Another decomposition is

$$(23468) = (23)(34)(46)(68).$$

266 Lemma Every permutation is the product of transpositions.

Proof: *It is enough to observe that*

$$(a_1\ a_2\ \ldots\ a_s) = (a_1\ a_s)(a_1\ a_{s-1}) \cdots (a_1\ a_2)$$

and appeal to Theorem 263. ❑

Let $\sigma \in S_n$ and let $(i, j) \in \{1, 2, \ldots, n\}^2$, $i \neq j$. Since σ is a permutation, $\exists (a, b) \in \{1, 2, \ldots, n\}^2$, $a \neq b$, such that $\sigma(j) - \sigma(i) = b - a$. This means that

$$\left| \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j} \right| = 1.$$

267 Definition Let $\sigma \in S_n$. We define the *sign* $\mathbf{sgn}(\sigma)$ of σ as

$$\mathbf{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(i) - \sigma(j)}{i - j} = (-1)^{\sigma}.$$

If $\mathbf{sgn}(\sigma) = 1$, then we say that σ is an *even permutation*, and if $\mathbf{sgn}(\sigma) = -1$ we say that σ is an *odd permutation*.

☞ *Notice that in fact*

$$\mathbf{sgn}(\sigma) = (-1)^{\mathbf{I}(\sigma)},$$

where $\mathbf{I}(\sigma) = \#\{(i, j) \mid 1 \leq i < j \leq n \textbf{ and } \sigma(i) > \sigma(j)\}$, *i.e.,* $\mathbf{I}(\sigma)$ *is the number of inversions that* σ *effects to the identity permutation* $\mathbf{Id}$.

268 Example The transposition $(1\ 2)$ has one inversion.

269 Lemma For any transposition $(k\ l)$ we have $\mathbf{sgn}((k\ l)) = -1$.

[1] A cycle of length 2 should more appropriately be called a *bicycle*.

Proof: *Let τ be transposition that exchanges k and l, and assume that $k < l$:*

$$\tau = \begin{bmatrix} 1 & 2 & \ldots & k-1 & k & k+1 & \ldots & l-1 & l & l+1 & \ldots & n \\ 1 & 2 & \ldots & k-1 & l & k+1 & \ldots & l-1 & k & l+1 & \ldots & n \end{bmatrix}$$

Let us count the number of inversions of τ:

- *The pairs (i, j) with $i \in \{1, 2, \ldots, k-1\} \cup \{l, l+1, \ldots, n\}$ and $i < j$ do not suffer an inversion;*
- *The pair (k, j) with $k < j$ suffers an inversion if and only if $j \in \{k+1, k+2, \ldots, l\}$, making $l - k$ inversions;*
- *If $i \in \{k+1, k+2, \ldots, l-1\}$ and $i < j$, (i, j) suffers an inversion if and only if $j = l$, giving $l - 1 - k$ inversions.*

This gives a total of $\mathbf{I}(\tau) = (l-k) + (l-1-k) = 2(l-k-1)+1$ inversions when $k < l$. Since this number is odd, we have $\mathbf{sgn}(\tau) = (-1)^{\mathbf{I}(\tau)} = -1$. In general we see that the transposition $(k\ l)$ has $2|k-l|-1$ inversions. ❑

270 Theorem Let $(\sigma, \tau) \in S_n^2$. Then

$$\mathbf{sgn}(\tau\sigma) = \mathbf{sgn}(\tau)\mathbf{sgn}(\sigma).$$

Proof: *We have*

$$\begin{aligned} \mathbf{sgn}(\sigma\tau) &= \prod_{1 \le i < j \le n} \frac{(\sigma\tau)(i) - (\sigma\tau)(j)}{i-j} \\ &= \left(\prod_{1 \le i < j \le n} \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)}\right) \cdot \left(\prod_{1 \le i < j \le n} \frac{\tau(i) - \tau(j)}{i-j}\right). \end{aligned}$$

The second factor on this last equality is clearly $\mathbf{sgn}(\tau)$, we must shew that the first factor is $\mathbf{sgn}(\sigma)$. Observe now that for $1 \le a < b \le n$ we have

$$\frac{\sigma(a) - \sigma(b)}{a-b} = \frac{\sigma(b) - \sigma(a)}{b-a}.$$

Since σ and τ are permutations, $\exists b \neq a$, $\tau(i) = a, \tau(j) = b$ and so $\sigma\tau(i) = \sigma(a), \sigma\tau(j) = b$. Thus

$$\frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)} = \frac{\sigma(a) - \sigma(b)}{a-b}$$

and so

$$\prod_{1 \le i < j \le n} \frac{\sigma(\tau(i)) - \sigma(\tau(j))}{\tau(i) - \tau(j)} = \prod_{1 \le a < b \le n} \frac{\sigma(a) - \sigma(b)}{a-b} = \mathbf{sgn}(\sigma).$$

❑

271 Corollary The identity permutation is even. If $\tau \in S_n$, then $\mathbf{sgn}(\tau) = \mathbf{sgn}(\tau^{-1})$.

Proof: *Since there are no inversions in $\mathbf{Id}$, we have $\mathbf{sgn}(\mathbf{Id}) = (-1)^0 = 1$. Since $\tau\tau^{-1} = \mathbf{Id}$, we must have $1 = \mathbf{sgn}(\mathbf{Id}) = \mathbf{sgn}(\tau\tau^{-1}) = \mathbf{sgn}(\tau)\mathbf{sgn}(\tau^{-1}) = (-1)^{\tau}(-1)^{\tau^{-1}}$ by Theorem 270. Since the values on the righthand of this last equality are ± 1, we must have $\mathbf{sgn}(\tau) = \mathbf{sgn}(\tau^{-1})$.* ❑

272 Lemma We have $\mathbf{sgn}(1\ 2\ \ldots\ l)) = (-1)^{l-1}$.

Proof: *Simply observe that the number of inversions of $(1\ 2\ \ldots\ l)$ is $l-1$.* ❑

273 Lemma Let $(\tau, (i_1\ \ldots\ i_l) \in S_n^2$. Then

$$\tau(i_1\ \ldots\ i_l)\tau^{-1} = (\tau(i_1)\ \ldots\ \tau(i_l)),$$

and if $\sigma \in S_n$ is a cycle of length l then

$$\mathbf{sgn}(\sigma) = (-1)^{l-1}$$

.

Proof: *For $1 \leq k < l$ we have $(\tau(i_1\ \ldots\ i_l)\tau^{-1})(\tau(i_k)) = \tau((i_1\ \ldots\ i_l)(i_k)) = \tau(i_{k+1})$. On a $(\tau(i_1\ \ldots\ i_l)\tau^{-1})(\tau(i_l)) = \tau((i_1\ \ldots\ i_l)(i_l)) = \tau(i_1)$. For $i \not\in \{\tau(i_1)\ \ldots\ \tau(i_l)\}$ we have $\tau^{-1}(i) \not\in \{i_1\ \ldots\ i_l\}$ whence $(i_1\ \ldots\ i_l)(\tau^{-1}(i)) = \tau^{-1}(i)$ etc.*

Furthermore, write $\sigma = (i_1\ \ldots\ i_l)$. Let $\tau \in S_n$ be such that $\tau(k) = i_k$ for $1 \leq k \leq l$. Then $\sigma = \tau(1\ 2\ \ldots\ l)\tau^{-1}$ and so we must have $\mathbf{sgn}(\sigma) = \mathbf{sgn}(\tau)\mathbf{sgn}((1\ 2\ \ldots\ l))\mathbf{sgn}(\tau^{-1})$, which equals $\mathbf{sgn}((1\ 2\ \ldots\ l))$ by virtue of Theorem 270 and Corollary 271. The result now follows by appealing to Lemma 272 ❑

274 Corollary Let $\sigma = \sigma_1\sigma_2\cdots\sigma_r$ be a product of disjoint cycles, each of length $l_1, \ldots, l_r$, respectively. Then

$$\mathbf{sgn}(\sigma) = (-1)^{\sum_{i=1}^{r}(l_i-1)}.$$

Hence, the product of two even permutations is even, the product of two odd permutations is even, and the product of an even permutation and an odd permutation is odd.

Proof: *This follows at once from Theorem 270 and Lemma 273.* ❑

275 Example The cycle (4678) is an odd cycle; the cycle (1) is an even cycle; the cycle (12345) is an even cycle.

276 Corollary Every permutation can be decomposed as a product of transpositions. This decomposition is not necessarily unique, but its parity is unique.

Proof: *This follows from Theorem 263, Lemma 266, and Corollary 274.* ❑

277 Example (The 15 puzzle) Consider a grid with 16 squares, as shewn in (6.1), where 15 squares are numbered 1 through 15 and the 16th slot is empty.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

(6.1)

In this grid we may successively exchange the empty slot with any of its neighbours, as for example

1	2	3	4
5	6	7	8
9	10	11	12
13	14		15

. (6.2)

We ask whether through a series of valid moves we may arrive at the following position.

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

(6.3)

Solution: ► *Let us shew that this is impossible. Each time we move a square to the empty position, we make transpositions on the set* $\{1, 2, \ldots, 16\}$*. Thus at each move, the permutation is multiplied by a transposition and hence it changes sign. Observe that the permutation corresponding to the square in (6.3) is* $(14\ 15)$ *(the positions 14th and 15th are transposed) and hence it is an odd permutation. But we claim that the empty slot can only return to its original position after an even permutation. To see this paint the grid as a checkerboard:*

B	R	B	R
R	B	R	B
B	R	B	R
R	B	R	B

(6.4)

We see that after each move, the empty square changes from black to red, and thus after an odd number of moves the empty slot is on a red square. Thus the empty slot cannot return to its original position in an odd number of moves. This completes the proof. ◄

Homework

Problem 6.2.1 Decompose the permutation

$$\begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 1 & 5 & 8 & 6 & 7 & 9 \end{bmatrix}$$

as a product of disjoint cycles and find its order.

6.3 Determinants

There are many ways of developing the theory of determinants. We will choose a way that will allow us to deduce the properties of determinants with ease, but has the drawback of being computationally cumbersome. In the next section we will shew that our way of defining determinants is equivalent to a more computationally friendly one.

It may be pertinent here to quickly review some properties of permutations. Recall that if $\sigma \in S_n$ is a cycle of length l, then its signum $\mathbf{sgn}(\sigma) = \pm 1$ depending on the parity of $l - 1$. For example, in S_7,

$$\sigma = (1\ 3\ 4\ 7\ 6)$$

has length 5, and the parity of $5 - 1 = 4$ is even, and so we write $\mathbf{sgn}(\sigma) = +1$. On the other hand,

$$\tau = (1\ 3\ 4\ 7\ 6\ 5)$$

has length 6, and the parity of $6-1=5$ is odd, and so we write $\mathbf{sgn}(\tau)=-1$.

Recall also that if $(\sigma,\tau)\in S_n^2$, then

$$\mathbf{sgn}(\tau\sigma)=\mathbf{sgn}(\tau)\mathbf{sgn}(\sigma).$$

Thus from the above two examples

$$\sigma\tau=(1\ 3\ 4\ 7\ 6)(1\ 3\ 4\ 7\ 6\ 5)$$

has signum $\mathbf{sgn}(\sigma)\mathbf{sgn}(\tau)=(+1)(-1)=-1$. Observe in particular that for the identity permutation $\mathbf{Id} \in S_n$ we have $\mathbf{sgn}(\mathbf{Id})=+1$.

278 Definition Let $A\in \mathbf{M}_{n\times n}(\mathbb{F}), A=[a_{ij}]$ be a square matrix. The *determinant of* A is defined and denoted by the sum

$$\mathbf{det}\, A=\sum_{\sigma\in S_n}\mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}.$$

☞ *The determinantal sum has* $n!$ *summands.*

279 Example If $n=1$, then S_1 has only one member, $\mathbf{Id}$, where $\mathbf{Id}(1)=1$. Since $\mathbf{Id}$ is an even permutation, $\mathbf{sgn}(\mathbf{Id})=(+1)$ Thus if $A=(a_{11})$, then

$$\mathbf{det}\, A=a_{11}$$

.

280 Example If $n=2$, then S_2 has $2!=2$ members, $\mathbf{Id}$ and $\sigma=(1\ 2)$. Observe that $\mathbf{sgn}(\sigma)=-1$. Thus if

$$A=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$$

then

$$\mathbf{det}\, A=\mathbf{sgn}(\mathbf{Id})a_{1\mathbf{Id}(1)}a_{2\mathbf{Id}(2)}+\mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}=a_{11}a_{22}-a_{12}a_{21}.$$

281 Example From the above formula for 2×2 matrices it follows that

$$\begin{aligned}\mathbf{det}\, A &= \mathbf{det}\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\\ &= (1)(4)-(3)(2)=-2,\\ \mathbf{det}\, B &= \mathbf{det}\begin{bmatrix}-1 & 2\\ 3 & 4\end{bmatrix}(-1)(4)-(3)(2)\\ &= -10,\end{aligned}$$

and

$$\mathbf{det}(A+B)=\mathbf{det}\begin{bmatrix}0 & 4\\ 6 & 8\end{bmatrix}=(0)(8)-(6)(4)=-24.$$

Observe in particular that $\mathbf{det}(A+B)\neq \mathbf{det}\, A+\mathbf{det}\, B$.

282 Example If $n = 3$, then S_2 has $3! = 6$ members:

$$\mathbf{Id}\ , \tau_1 = (2\ 3), \tau_2 = (1\ 3), \tau_3 = (1\ 2), \sigma_1 = (1\ 2\ 3), \sigma_2 = (1\ 3\ 2).$$

. Observe that $\mathbf{Id}$, σ_1, σ_2 are even, and τ_1, τ_2, τ_3 are odd. Thus if

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

then

$$\begin{aligned} \mathbf{det}\, A &= \mathbf{sgn}(\mathbf{Id}\ )a_{1\mathbf{Id}\ (1)}a_{2\mathbf{Id}\ (2)}a_{3\mathbf{Id}\ (3)} + \mathbf{sgn}(\tau_1)a_{1\tau_1(1)}a_{2\tau_1(2)}a_{3\tau_1(3)} \\ &\quad + \mathbf{sgn}(\tau_2)a_{1\tau_2(1)}a_{2\tau_2(2)}a_{3\tau_2(3)} + \mathbf{sgn}(\tau_3)a_{1\tau_3(1)}a_{2\tau_3(2)}a_{3\tau_3(3)} \\ &\quad + \mathbf{sgn}(\sigma_1)a_{1\sigma_1(1)}a_{2\sigma_1(2)}a_{3\sigma_1(3)} + \mathbf{sgn}(\sigma_2)a_{1\sigma_2(1)}a_{2\sigma_2(2)}a_{3\sigma_2(3)} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{13}a_{22}a_{31} \\ &\quad - a_{13}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}. \end{aligned}$$

283 Theorem (Row-Alternancy of Determinants) Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$. If $B \in \mathbf{M}_{n\times n}(\mathbb{F}), B = [b_{ij}]$ is the matrix obtained by interchanging the s-th row of A with its t-th row, then $\mathbf{det}\, B = -\,\mathbf{det}\, A$.

Proof: *Let τ be the transposition*

$$\tau = \begin{bmatrix} s & t \\ \tau(t) & \tau(s) \end{bmatrix}.$$

Then $\sigma\tau(a) = \sigma(a)$ for $a \in \{1, 2, \ldots, n\} \setminus \{s, t\}$. Also, $\mathbf{sgn}(\sigma\tau) = \mathbf{sgn}(\sigma)\mathbf{sgn}(\tau) = -\mathbf{sgn}(\sigma)$. As σ ranges through all permutations of S_n, so does $\sigma\tau$, hence

$$\begin{aligned} \mathbf{det}\, B &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma)b_{1\sigma(1)}b_{2\sigma(2)}\cdots b_{s\sigma(s)}\cdots b_{t\sigma(t)}\cdots b_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{st}\cdots a_{ts}\cdots a_{n\sigma(n)} \\ &= -\sum_{\sigma\in S_n} \mathbf{sgn}(\sigma\tau)a_{1\sigma\tau(1)}a_{2\sigma\tau(2)}\cdots a_{s\sigma\tau(s)}\cdots a_{t\sigma\tau(t)}\cdots a_{n\sigma\tau(n)} \\ &= -\sum_{\lambda\in S_n} \mathbf{sgn}(\lambda)a_{1\lambda(1)}a_{2\lambda(2)}\cdots a_{n\lambda(n)} \\ &= -\,\mathbf{det}\, A. \end{aligned}$$

❑

284 Corollary If $A_{(r:k)}, 1 \leq k \leq n$ denote the rows of A and $\sigma \in S_n$, then

$$\mathbf{det}\begin{bmatrix} A_{(r:\sigma(1))} \\ A_{(r:\sigma(2))} \\ \vdots \\ A_{(r:\sigma(n))} \end{bmatrix} = (\mathbf{sgn}(\sigma))\,\mathbf{det}\,A.$$

An analogous result holds for columns.

Proof: *Apply the result of Theorem 283 multiple times.* ❑

285 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$. Then

$$\mathbf{det}\,A^T = \mathbf{det}\,A.$$

Proof: *Let* $C = A^T$. *By definition*

$$\begin{aligned} \mathbf{det}\,A^T &= \mathbf{det}\,C \\ &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma)c_{1\sigma(1)}c_{2\sigma(2)}\cdots c_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma)a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}. \end{aligned}$$

But the product $a_{\sigma(1)1}a_{\sigma(2)2}\cdots a_{\sigma(n)n}$ *also appears in* $\mathbf{det}\,A$ *with the same signum* $\mathbf{sgn}(\sigma)$, *since the permutation*

$$\begin{bmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{bmatrix}$$

is the inverse of the permutation

$$\begin{bmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{bmatrix}.$$

❑

286 Corollary (Column-Alternancy of Determinants) Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$. If $C \in \mathbf{M}_{n\times n}(\mathbb{F}), C = [c_{ij}]$ is the matrix obtained by interchanging the s-th column of A with its t-th column, then $\mathbf{det}\,C = -\mathbf{det}\,A$.

Proof: *This follows upon combining Theorem 283 and Theorem 285.* ❑

287 Theorem (Row Homogeneity of Determinants) Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$ and $\alpha \in \mathbb{F}$. If $B \in \mathbf{M}_{n\times n}(\mathbb{F}), B = [b_{ij}]$ is the matrix obtained by multiplying the s-th row of A by α, then

$$\mathbf{det}\,B = \alpha\,\mathbf{det}\,A.$$

Proof: *Simply observe that*

$$\mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots \alpha a_{s\sigma(s)}\cdots a_{n\sigma(n)} = \alpha\mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{s\sigma(s)}\cdots a_{n\sigma(n)}.$$

❑

288 Corollary (Column Homogeneity of Determinants) If $C \in \mathbf{M}_{n\times n}(\mathbb{F}), C = (C_{ij})$ is the matrix obtained by multiplying the s-th column of A by α, then

$$\mathbf{det}\, C = \alpha\, \mathbf{det}\, A.$$

Proof: *This follows upon using Theorem 285 and Theorem 287.* ❑

☞ *It follows from Theorem 287 and Corollary 288 that if a row (or column) of a matrix consists of $0_{\mathbb{F}}$s only, then the determinant of this matrix is $0_{\mathbb{F}}$.*

289 Example

$$\mathbf{det}\begin{bmatrix} x & 1 & a \\ x^2 & 1 & b \\ x^3 & 1 & c \end{bmatrix} = x\,\mathbf{det}\begin{bmatrix} 1 & 1 & a \\ x & 1 & b \\ x^2 & 1 & c \end{bmatrix}.$$

290 Corollary

$$\mathbf{det}(\alpha A) = \alpha^n\, \mathbf{det}\, A.$$

Proof: *Since there are n columns, we are able to pull out one factor of α from each one.* ❑

291 Example Recall that a matrix A is *skew-symmetric* if $A = -A^T$. Let $A \in \mathbf{M}_{2001}(\mathbb{R})$ be skew-symmetric. Prove that $\mathbf{det}\, A = 0$.

Solution: ► *We have*

$$\mathbf{det}\, A = \mathbf{det}(-A^T) = (-1)^{2001}\, \mathbf{det}\, A^T = -\,\mathbf{det}\, A,$$

and so $2\,\mathbf{det}\, A = 0$, *from where* $\mathbf{det}\, A = 0$. ◄

292 Lemma (Row-Linearity and Column-Linearity of Determinants) Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$. For a fixed row s, suppose that $a_{sj} = b_{sj} + c_{sj}$ for each $j \in [1; n]$. Then

$$\mathbf{det}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{(s-1)1} & a_{(s-1)2} & \cdots & a_{(s-1)n} \\ b_{s1}+c_{s1} & b_{s2}+c_{s2} & \cdots & b_{sn}+c_{sn} \\ a_{(s+1)1} & a_{(s+1)2} & \cdots & a_{(s+1)n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$= \mathbf{det}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{(s-1)1} & a_{(s-1)2} & \cdots & a_{(s-1)n} \\ b_{s1} & b_{s2} & \cdots & b_{sn} \\ a_{(s+1)1} & a_{(s+1)2} & \cdots & a_{(s+1)n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$+\,\mathbf{det}\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{(s-1)1} & a_{(s-1)2} & \cdots & a_{(s-1)n} \\ c_{s1} & c_{s2} & \cdots & c_{sn} \\ a_{(s+1)1} & a_{(s+1)2} & \cdots & a_{(s+1)n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

An analogous result holds for columns.

Proof: *Put*

$$S = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{(s-1)1} & a_{(s-1)2} & \cdots & a_{(s-1)n} \\ b_{s1}+c_{s1} & b_{s2}+c_{s2} & \cdots & b_{sn}+c_{sn} \\ a_{(s+1)1} & a_{(s+1)2} & \cdots & a_{(s+1)n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix},$$

$$T = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{(s-1)1} & a_{(s-1)2} & \cdots & a_{(s-1)n} \\ b_{s1} & b_{s2} & \cdots & b_{sn} \\ a_{(s+1)1} & a_{(s+1)2} & \cdots & a_{(s+1)n} \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

and

$$U = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{(s-1)1} & a_{(s-1)2} & \cdots & a_{(s-1)n} \\ c_{s1} & c_{s2} & \cdots & c_{sn} \\ a_{(s+1)1} & a_{(s+1)2} & \cdots & a_{(s+1)n} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}.$$

Then

$$\begin{aligned} \mathbf{det}\, S &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{(s-1)\sigma(s-1)} (b_{s\sigma(s)} \\ &\qquad + c_{s\sigma(s)}) a_{(s+1)\sigma(s+1)} \cdots a_{n\sigma(n)} \\ &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{(s-1)\sigma(s-1)} b_{s\sigma(s)} a_{(s+1)\sigma(s+1)} \cdots a_{n\sigma(n)} \\ &\quad + \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{(s-1)\sigma(s-1)} c_{s\sigma(s)} a_{(s+1)\sigma(s+1)} \cdots a_{n\sigma(n)} \\ &= \mathbf{det}\, T + \mathbf{det}\, U. \end{aligned}$$

By applying the above argument to $\mathbf{A}^T$*, we obtain the result for columns.*
❑

293 Lemma If two rows or two columns of $\mathbf{A} \in \mathbf{M}_{n\times n}(\mathbb{F})$, $\mathbf{A} = [a_{ij}]$ are identical, then $\mathbf{det}\, \mathbf{A} = 0_{\mathbb{F}}$.

Proof: *Suppose* $a_{sj} = a_{tj}$ *for* $s \neq t$ *and for all* $j \in [1; n]$*. In particular, this means that for any* $\sigma \in S_n$ *we have* $a_{s\sigma(t)} = a_{t\sigma(t)}$ *and* $a_{t\sigma(s)} = a_{s\sigma(s)}$*. Let* τ *be the transposition*

$$\tau = \begin{bmatrix} s & t \\ \tau(t) & \tau(s) \end{bmatrix}.$$

Then $\sigma\tau(a) = \sigma(a)$ *for* $a \in \{1, 2, \ldots, n\} \setminus \{s, t\}$*. Also,* $\mathbf{sgn}(\sigma\tau) = \mathbf{sgn}(\sigma)\mathbf{sgn}(\tau) = -\mathbf{sgn}(\sigma)$*. As* σ *runs through all even permutations,* $\sigma\tau$ *runs through all odd permutations, and viceversa. Therefore*

$$\begin{aligned} \det A &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{s\sigma(s)} \cdots a_{t\sigma(t)} \cdots a_{n\sigma(n)} \\ &= \sum_{\substack{\sigma\in S_n \\ \mathbf{sgn}(\sigma)=1}} \big(\mathbf{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{s\sigma(s)} \cdots a_{t\sigma(t)} \cdots a_{n\sigma(n)} \\ &\qquad + \mathbf{sgn}(\sigma\tau) a_{1\sigma\tau(1)} a_{2\sigma\tau(2)} \cdots a_{s\sigma\tau(s)} \cdots a_{t\sigma\tau(t)} \cdots a_{n\sigma\tau(n)}\big) \\ &= \sum_{\substack{\sigma\in S_n \\ \mathbf{sgn}(\sigma)=1}} \mathbf{sgn}(\sigma) \big(a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{s\sigma(s)} \cdots a_{t\sigma(t)} \cdots a_{n\sigma(n)} \\ &\qquad - a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{s\sigma(t)} \cdots a_{t\sigma(s)} \cdots a_{n\sigma(n)}\big) \\ &= \sum_{\substack{\sigma\in S_n \\ \mathbf{sgn}(\sigma)=1}} \mathbf{sgn}(\sigma) \big(a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{s\sigma(s)} \cdots a_{t\sigma(t)} \cdots a_{n\sigma(n)} \\ &\qquad - a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{t\sigma(t)} \cdots a_{s\sigma(s)} \cdots a_{n\sigma(n)}\big) \\ &= 0_{\mathbb{F}}. \end{aligned}$$

Arguing on $\mathbf{A}^T$ *will yield the analogous result for the columns.* ❑

294 Corollary If two rows or two columns of $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$ are proportional, then $\mathbf{det}\, A = 0_{\mathbb{F}}$.

Proof: *Suppose* $a_{sj} = \alpha a_{tj}$ *for* $s \neq t$ *and for all* $j \in [1; n]$. *If* B *is the matrix obtained by pulling out the factor* α *from the* s*-th row then* $\mathbf{det}\, A = \alpha\, \mathbf{det}\, B$. *But now the* s*-th and the* t*-th rows in* B *are identical, and so* $\mathbf{det}\, B = 0_{\mathbb{F}}$. *Arguing on* A^T *will yield the analogous result for the columns.*

❑

295 Example

$$\mathbf{det}\begin{bmatrix} 1 & a & b \\ 1 & a & c \\ 1 & a & d \end{bmatrix} = a\,\mathbf{det}\begin{bmatrix} 1 & 1 & b \\ 1 & 1 & c \\ 1 & 1 & d \end{bmatrix} = 0,$$

since on the last determinant the first two columns are identical.

296 Theorem (Multilinearity of Determinants) Let $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$ and $\alpha \in \mathbb{F}$. If $X \in \mathbf{M}_{n\times n}(\mathbb{F}), X = (x_{ij})$ is the matrix obtained by the row transvection $R_s + \alpha R_t \to R_s$ then $\mathbf{det}\, X = \mathbf{det}\, A$. Similarly, if $Y \in \mathbf{M}_{n\times n}(\mathbb{F}), Y = (y_{ij})$ is the matrix obtained by the column transvection $C_s + \alpha C_t \to C_s$ then $\mathbf{det}\, Y = \mathbf{det}\, A$.

Proof: *For the row transvection it suffices to take* $b_{sj} = a_{sj}$, $c_{sj} = \alpha a_{tj}$ *for* $j \in [1; n]$ *in Lemma 292. With the same notation as in the lemma,* $T = A$, *and so,*

$$\mathbf{det}\, X = \mathbf{det}\, T + \mathbf{det}\, U = \mathbf{det}\, A + \mathbf{det}\, U.$$

But U *has its* s*-th and* t*-th rows proportional* ($s \neq t$), *and so by Corollary 294* $\mathbf{det}\, U = 0_{\mathbb{F}}$. *Hence* $\mathbf{det}\, X = \mathbf{det}\, A$. *To obtain the result for column transvections it suffices now to also apply Theorem 285.* ❑

297 Example Demonstrate, *without actually calculating the determinant* that

$$\mathbf{det}\begin{bmatrix} 2 & 9 & 9 \\ 4 & 6 & 8 \\ 7 & 4 & 1 \end{bmatrix}$$

is divisible by 13.

Solution: ► *Observe that* 299, 468 *and* 741 *are all divisible by 13. Thus*

$$\mathbf{det}\begin{bmatrix} 2 & 9 & 9 \\ 4 & 6 & 8 \\ 7 & 4 & 1 \end{bmatrix} \overset{C_3+10C_2+100C_1\to C_3}{=} \mathbf{det}\begin{bmatrix} 2 & 9 & 299 \\ 4 & 6 & 468 \\ 7 & 4 & 741 \end{bmatrix} = 13\,\mathbf{det}\begin{bmatrix} 2 & 9 & 23 \\ 4 & 6 & 36 \\ 7 & 4 & 57 \end{bmatrix},$$

which shews that the determinant is divisible by 13. ◄

298 Theorem The determinant of a triangular matrix (upper or lower) is the product of its diagonal elements.

Proof: *Let* $A \in \mathbf{M}_{n\times n}(\mathbb{F}), A = [a_{ij}]$ *be a triangular matrix. Observe that if* $\sigma \neq \mathbf{Id}$ *then* $a_{i\sigma(i)}a_{\sigma(i)\sigma^2(i)} = 0_{\mathbb{F}}$ *occurs in the product*

$$a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}.$$

Thus

$$\begin{aligned}\mathbf{det}\, A &= \sum_{\sigma\in S_n} \mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)} \\ &= \mathbf{sgn}(\mathbf{Id}\,)a_{1\mathbf{Id}\,(1)}a_{2\mathbf{Id}\,(2)}\cdots a_{n\mathbf{Id}\,(n)} = a_{11}a_{22}\cdots a_{nn}.\end{aligned}$$

❑

299 Example The determinant of the $n \times n$ identity matrix $\mathbf{I}_n$ over a field $\mathbb{F}$ is

$$\mathbf{det}\, \mathbf{I}_n = 1_{\mathbb{F}}.$$

300 Example Find

$$\mathbf{det} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}.$$

Solution: ► *We have*

$$\begin{aligned}\mathbf{det} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} &\overset{\substack{C_2-2C_1\to C_2 \\ C_3-3C_1\to C_3}}{\rightsquigarrow} \mathbf{det} \begin{bmatrix} 1 & 0 & 0 \\ 4 & -3 & -6 \\ 7 & -6 & -12 \end{bmatrix} \\ &= (-3)(-6)\,\mathbf{det} \begin{bmatrix} 1 & 0 & 0 \\ 4 & 1 & 1 \\ 7 & 2 & 2 \end{bmatrix} \\ &= 0,\end{aligned}$$

since in this last matrix the second and third columns are identical and so Lemma 293 applies. ◄

301 Theorem Let $(A, B) \in (\mathbf{M}_{n\times n}(\mathbb{F}))^2$. Then

$$\mathbf{det}(AB) = (\mathbf{det}\, A)(\mathbf{det}\, B).$$

Proof: *Put* $D = AB, D = (d_{ij}), d_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. *If* $A_{(c:k)}, D_{(c:k)}, 1 \leq k \leq n$ *denote the columns of* A *and* D, *respectively, observe that*

$$D_{(c:k)} = \sum_{l=1}^{n} b_{lk}A_{(c:l)}, \quad 1 \leq k \leq n.$$

Applying Corollary 288 and Lemma 292 multiple times, we obtain

$$\begin{aligned}\det D &= \det(D_{(c:1)}, D_{(c:2)}, \ldots, D_{(c:n)}) \\ &= \sum_{j_1=1}^{n} \sum_{j_2=1}^{n} \cdots \sum_{j_n=1}^{n} b_{1j_1} b_{2j_2} \cdots b_{nj_n} \\ &\quad \cdot \det(A_{(c:j_1)}, A_{(c:j_2)}, \ldots, A_{(c:j_n)}).\end{aligned}$$

By Lemma 293, if any two of the $A_{(c:j_l)}$ are identical, the determinant on the right vanishes. So each one of the j_l is different in the non-vanishing terms and so the map

$$\sigma : \begin{array}{ccc} \{1, 2, \ldots, n\} & \to & \{1, 2, \ldots, n\} \\ l & \mapsto & j_l \end{array}$$

is a permutation. Here $j_l = \sigma(l)$. Therefore, for the non-vanishing

$$\det(A_{(c:j_1)}, A_{(c:j_2)}, \ldots, A_{(c:j_n)})$$

we have in view of Corollary 284,

$$\begin{aligned}\det(A_{(c:j_1)}, A_{(c:j_2)}, \ldots, A_{(c:j_n)}) &= (\mathbf{sgn}(\sigma)) \det(A_{(c:1)}, A_{(c:2)}, \ldots, A_{(c:n)}) \\ &= (\mathbf{sgn}(\sigma)) \det A.\end{aligned}$$

We deduce that

$$\begin{aligned}\det(AB) &= \det D \\ &= \sum_{j_n=1}^{n} b_{1j_1} b_{2j_2} \cdots b_{nj_n} \det(A_{(c:j_1)}, A_{(c:j_2)}, \ldots, A_{(c:j_n)}) \\ &= (\det A) \sum_{\sigma \in S_n} (\mathbf{sgn}(\sigma)) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)} \\ &= (\det A)(\det B),\end{aligned}$$

as we wanted to shew. ❑

By applying the preceding theorem multiple times we obtain

302 Corollary If $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ and if k is a positive integer then

$$\det A^k = (\det A)^k.$$

303 Corollary If $A \in \mathbf{GL}_n(\mathbb{F})$ and if k is a positive integer then $\det A \neq 0_{\mathbb{F}}$ and

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

Proof: *We have $AA^{-1} = \mathbf{I}_n$ and so by Theorem 301 $(\det A)(\det A^{-1}) = 1_{\mathbb{F}}$, from where the result follows.* ❑

Homework

Problem 6.3.1 Let

$$\Omega = \det \begin{bmatrix} bc & ca & ab \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}.$$

Without expanding either determinant, prove that

$$\Omega = \mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{bmatrix}.$$

Problem 6.3.2 Demonstrate that

$$\Omega = \mathbf{det}\begin{bmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{bmatrix} = (a+b+c)^3.$$

Problem 6.3.3 After the indicated column operations on a 3×3 matrix A with $\mathbf{det}\, A = -540$, matrices $A_1, A_2, \ldots, A_5$ are successively obtained:

$$A \overset{C_1+3C_2\to C_1}{\to} A_1 \overset{C_2\leftrightarrow C_3}{\to} A_2 \overset{3C_2-C_1\to C_2}{\to} A_3 \overset{C_1-3C_2\to C_1}{\to} A_4 \overset{2C_1\to C_1}{\to} A_5$$

Determine the numerical values of $\mathbf{det}\, A_1$, $\mathbf{det}\, A_2$, $\mathbf{det}\, A_3$, $\mathbf{det}\, A_4$ and $\mathbf{det}\, A_5$.

Problem 6.3.4 Prove, without actually expanding the determinant, that

$$\mathbf{det}\begin{bmatrix} 1 & 2 & 3 & 7 & 0 \\ 6 & 1 & 5 & 14 & 1 \\ 8 & 6 & 1 & 21 & 3 \\ 7 & 3 & 8 & 7 & 1 \\ 2 & 4 & 6 & 0 & 4 \end{bmatrix}$$

is divisible by 1722.

Problem 6.3.5 Let A, B, C be 3×3 matrices with $\mathbf{det}\, A = 3, \mathbf{det}\, B^3 = -8, \mathbf{det}\, C = 2$. Compute (i) $\mathbf{det}\, ABC$, (ii) $\mathbf{det}\, 5AC$, (iii) $\mathbf{det}\, A^3B^{-3}C^{-1}$. Express your answers as fractions.

Problem 6.3.6 Shew that $\forall A \in \mathbf{M}_{n\times n}(\mathbb{R})$,

$$\exists (X, Y) \in (\mathbf{M}_{n\times n}(\mathbb{R}))^2, (\mathbf{det}\, X)(\mathbf{det}\, Y) \neq 0$$

such that

$$A = X + Y.$$

That is, any square matrix over $\mathbb{R}$ can be written as a sum of two matrices whose determinant is not zero.

Problem 6.3.7 Prove or disprove! The set $X = \{A \in \mathbf{M}_{n\times n}(\mathbb{F}) : \mathbf{det}\, A = 0_{\mathbb{F}}\}$ is a vector subspace of $\mathbf{M}_{n\times n}(\mathbb{F})$.

6.4 Laplace Expansion

We now develop a more computationally convenient approach to determinants.

Put

$$C_{ij} = \sum_{\substack{\sigma\in S_n \\ \sigma(i)=j}} (\mathbf{sgn}(\sigma))a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}.$$

Then

$$\begin{aligned}\mathbf{det}\, A &= \sum_{\sigma\in S_n}(\mathbf{sgn}(\sigma))a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}\\ &= \sum_{i=1}^{n} a_{ij}\sum_{\substack{\sigma\in S_n\\ \sigma(i)=j}}(\mathbf{sgn}(\sigma))a_{1\sigma(1)}a_{2\sigma(2)}\\ &\qquad\cdots a_{(i-1)\sigma(i-1)}a_{(i+1)\sigma(i+1)}\cdots a_{n\sigma(n)}\\ &= \sum_{i=1}^{n} a_{ij}C_{ij},\end{aligned} \tag{6.5}$$

is the expansion of $\mathbf{det}\, A$ along the j-th column. Similarly,

$$\begin{aligned}\mathbf{det}\, A &= \sum_{\sigma\in S_n}(\mathbf{sgn}(\sigma))a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}\\ &= \sum_{j=1}^{n} a_{ij}\sum_{\substack{\sigma\in S_n\\ \sigma(i)=j}}(\mathbf{sgn}(\sigma))a_{1\sigma(1)}a_{2\sigma(2)}\\ &\qquad\cdots a_{(i-1)\sigma(i-1)}a_{(i+1)\sigma(i+1)}\cdots a_{n\sigma(n)}\\ &= \sum_{j=1}^{n} a_{ij}C_{ij},\end{aligned}$$

is the expansion of $\mathbf{det}\, A$ along the i-th row.

304 Definition Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$, $A = [a_{ij}]$. The ij-th minor $A_{ij} \in \mathbf{M}_{n-1}(\mathbb{R})$ is the $(n-1)\times(n-1)$ matrix obtained by deleting the i-th row and the j-th column from A.

305 Example If

$$A = \begin{bmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 7 & 8 & 9\end{bmatrix}$$

then, for example,

$$A_{11} = \begin{bmatrix}5 & 6\\ 8 & 9\end{bmatrix},\quad A_{12} = \begin{bmatrix}4 & 6\\ 7 & 9\end{bmatrix},\quad A_{21} = \begin{bmatrix}2 & 3\\ 8 & 9\end{bmatrix},\quad A_{22} = \begin{bmatrix}1 & 3\\ 7 & 9\end{bmatrix},\quad A_{33} = \begin{bmatrix}1 & 2\\ 4 & 5\end{bmatrix}.$$

306 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then

$$\mathbf{det}\, A = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\,\mathbf{det}\, A_{ij} = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}\,\mathbf{det}\, A_{ij}.$$

Proof: *It is enough to shew, in view of 6.5 that*

$$(-1)^{i+j}\,\mathbf{det}\, A_{ij} = C_{ij}.$$

Now,

$$\begin{aligned}C_{nn} &= \sum_{\substack{\sigma\in S_n\\ \sigma(n)=n}}\mathbf{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{(n-1)\sigma(n-1)}\\ &= \sum_{\tau\in S_{n-1}}\mathbf{sgn}(\tau)a_{1\tau(1)}a_{2\tau(2)}\cdots a_{(n-1)\tau(n-1)}\\ &= \mathbf{det}\, A_{nn},\end{aligned}$$

since the second sum shewn is the determinant of the submatrix obtained by deleting the last row and last column from A.

To find C_{ij} *for general* ij *we perform some row and column interchanges to* A *in order to bring* a_{ij} *to the* nn*-th position. We thus bring the* i*-th row to the* n*-th row by a series of transpositions, first swapping the* i*-th and the* $(i+1)$*-th row, then swapping the new* $(i+1)$*-th row and the* $(i+2)$*-th row, and so forth until the original* i*-th row makes it to the* n*-th row. We have made thereby* $n-i$ *interchanges. To this new matrix we perform analogous interchanges to the* j*-th column, thereby making* $n-j$ *interchanges. We have made a total of* $2n-i-j$ *interchanges. Observe that* $(-1)^{2n-i-j} = (-1)^{i+j}$. *Call the analogous quantities in the resulting matrix* A', C'_{nn}, A'_{nn}. *Then*

$$C_{ij} = C'_{nn} = \mathbf{det}\, A'_{nn} = (-1)^{i+j}\, \mathbf{det}\, A_{ij},$$

by virtue of Corollary 284.

❑

☞ *It is irrelevant which row or column we choose to expand a determinant of a square matrix. We always obtain the same result. The sign pattern is given by*

$$\begin{bmatrix} + & - & + & - & \cdots \\ - & + & - & + & \vdots \\ + & - & + & - & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}$$

307 Example Find

$$\mathbf{det}\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

by expanding along the first row.

Solution: ▶ *We have*

$$\begin{aligned} \mathbf{det}\, A &= 1(-1)^{1+1}\,\mathbf{det}\begin{bmatrix} 5 & 6 \\ 8 & 9 \end{bmatrix} + 2(-1)^{1+2}\,\mathbf{det}\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 3(-1)^{1+3}\,\mathbf{det}\begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix} \\ &= 1(45-48) - 2(36-42) + 3(32-35) = 0. \end{aligned}$$

◀

308 Example Evaluate the *Vandermonde* determinant

$$\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}.$$

Solution: ▶

$$\det\begin{bmatrix}1 & 1 & 1\\ a & b & c\\ a^2 & b^2 & c^2\end{bmatrix} = \det\begin{bmatrix}1 & 0 & 0\\ a & b-a & c-a\\ a^2 & b^2-a^2 & c^2-a^2\end{bmatrix}$$

$$= \det\begin{bmatrix}b-a & c-a\\ b^2-c^2 & c^2-a^2\end{bmatrix}$$

$$= (b-a)(c-a)\det\begin{bmatrix}1 & 1\\ b+a & c+a\end{bmatrix}$$

$$= (b-a)(c-a)(c-b).$$

◀

309 Example Evaluate the determinant

$$\det A = \det\begin{bmatrix}1 & 2 & 3 & 4 & \cdots & 2000\\ 2 & 1 & 2 & 3 & \cdots & 1999\\ 3 & 2 & 1 & 2 & \cdots & 1998\\ 4 & 3 & 2 & 1 & \cdots & 1997\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ 2000 & 1999 & 1998 & 1997 & \cdots & 1\end{bmatrix}.$$

Solution: ▶ *Applying* $R_n - R_{n+1} \to R_n$ *for* $1 \leq n \leq 1999$, *the determinant becomes*

$$\det\begin{bmatrix}-1 & 1 & 1 & 1 & \cdots & 1 & 1\\ -1 & -1 & 1 & 1 & \cdots & 1 & 1\\ -1 & -1 & -1 & 1 & \cdots & 1 & 1\\ -1 & -1 & -1 & -1 & \cdots & 1 & 1\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\ -1 & -1 & -1 & -1 & \cdots & -1 & 1\\ 2000 & 1999 & 1998 & 1997 & \cdots & 2 & 1\end{bmatrix}.$$

Applying now $C_n + C_{2000} \to C_n$ *for* $1 \leq n \leq 1999$, *we obtain*

$$\mathbf{det}\begin{bmatrix} 0 & 2 & 2 & 2 & \cdots & 2 & 1 \\ 0 & 0 & 2 & 2 & \cdots & 2 & 1 \\ 0 & 0 & 0 & 2 & \cdots & 2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 2 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 2001 & 2000 & 1999 & 1998 & \cdots & 3 & 1 \end{bmatrix}.$$

This last determinant we expand along the first column. We have

$$2001\,\mathbf{det}\begin{bmatrix} 2 & 2 & 2 & \cdots & 2 & 1 \\ 0 & 2 & 2 & \cdots & 2 & 1 \\ 0 & 0 & 2 & \cdots & 2 & 1 \\ 0 & 0 & 0 & \cdots & 2 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix} = 2001(2^{1998}).$$

◀

310 Definition Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. The *classical adjoint* or *adjugate* of A is the $n \times n$ matrix $\mathbf{adj}\,(A)$ whose entries are given by

$$[\mathbf{adj}\,(A)]_{ij} = (-1)^{i+j}\,\mathbf{det}\,A_{ji},$$

where A_{ji} is the ji-th minor of A.

311 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then

$$(\mathbf{adj}\,(A))A = A(\mathbf{adj}\,(A)) = (\mathbf{det}\,A)\mathbf{I}_n.$$

Proof: *We have*

$$\begin{aligned} [A(\mathbf{adj}\,(A))]_{ij} &= \sum_{k=1}^{n} a_{ik}[\mathbf{adj}\,(A)]_{kj} \\ &= \sum_{k=1}^{n} a_{ik}(-1)^{i+k}\,\mathbf{det}\,A_{jk}. \end{aligned}$$

Now, this last sum is $\mathbf{det}\,A$ *if* $i = j$ *by virtue of Theorem 306. If* $i \neq j$ *it is* 0, *since then the* j*-th row is identical to the* i*-th row and this determinant is* $0_{\mathbb{F}}$ *by virtue of Lemma 293. Thus on the diagonal entries we get* $\mathbf{det}\,A$ *and the off-diagonal entries are* $0_{\mathbb{F}}$. *This proves the theorem.* ❑

The next corollary follows immediately.

312 Corollary Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then A is invertible if and only $\mathbf{det}\,A \neq 0_{\mathbb{F}}$ and

$$A^{-1} = \frac{\mathbf{adj}\,(A)}{\mathbf{det}\,A}.$$

Homework

Problem 6.4.1 Find

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

by expanding along the second column.

Problem 6.4.2 Prove that $\det \begin{bmatrix} a & b & c \\ c & a & b \\ b & c & a \end{bmatrix} = a^3 + b^3 + c^3 - 3abc$. This type of matrix is called a *circulant* matrix.

Problem 6.4.3 Compute the determinant

$$\det \begin{bmatrix} 1 & 0 & -1 & 1 \\ 2 & 0 & 0 & 1 \\ 666 & -3 & -1 & 1000000 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Problem 6.4.4 Prove that

$$\det \begin{bmatrix} x+a & b & c \\ a & x+b & c \\ a & b & x+c \end{bmatrix} = x^2(x+a+b+c).$$

Problem 6.4.5 If

$$\det \begin{bmatrix} 1 & 1 & 1 & 1 \\ x & a & 0 & 0 \\ x & 0 & b & 0 \\ x & 0 & 0 & c \end{bmatrix} = 0,$$

and $xabc \neq 0$, prove that

$$\frac{1}{x} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}.$$

Problem 6.4.6 Consider the matrix

$$A = \begin{bmatrix} a & -b & -c & -d \\ b & a & d & -c \\ c & -d & a & b \\ d & c & -b & a \end{bmatrix}.$$

❶ Compute $A^T A$.

❷ Use the above to prove that

$$\mathbf{det}\, A = (a^2 + b^2 + c^2 + d^2)^2.$$

Problem 6.4.7 Prove that

$$\mathbf{det} \begin{bmatrix} 0 & a & b & 0 \\ a & 0 & b & 0 \\ 0 & a & 0 & b \\ 1 & 1 & 1 & 1 \end{bmatrix} = 2ab(a-b).$$

Problem 6.4.8 Demonstrate that

$$\mathbf{det} \begin{bmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{bmatrix} = (ad - bc)^2.$$

Problem 6.4.9 Use induction to shew that

$$\mathbf{det} \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} = (-1)^{n+1}.$$

Problem 6.4.10 Let

$$A = \begin{bmatrix} 1 & n & n & n & \cdots & n \\ n & 2 & n & n & \vdots & n \\ n & n & 3 & n & \cdots & n \\ n & n & n & 4 & \cdots & n \\ \vdots & \vdots & \vdots & \cdots & \vdots & \\ n & n & n & n & n & n \end{bmatrix},$$

that is, $A \in \mathbf{M}_{n\times n}(\mathbb{R})$, $A = [a_{ij}]$ is a matrix such that $a_{kk} = k$ and $a_{ij} = n$ when $i \neq j$. Find $\mathbf{det}\, A$.

Problem 6.4.11 Let $n \in \mathbb{N}$, $n > 1$ be an odd integer. Recall that the binomial coefficients $\binom{n}{k}$ satisfy $\binom{n}{n} = \binom{n}{0} = 1$ and that for $1 \leq k \leq n$,

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}.$$

Prove that

$$\det\begin{bmatrix} 1 & \binom{n}{1} & \binom{n}{2} & \cdots & \binom{n}{n-1} & 1 \\ 1 & 1 & \binom{n}{1} & \cdots & \binom{n}{n-2} & \binom{n}{n-1} \\ \binom{n}{n-1} & 1 & 1 & \cdots & \binom{n}{n-3} & \binom{n}{n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \binom{n}{1} & \binom{n}{2} & \binom{n}{3} & \cdots & 1 & 1 \end{bmatrix} = (1+(-1)^n)^n.$$

Problem 6.4.12 Let $A \in \mathbf{GL}_n(\mathbb{F})$, $n > 1$. Prove that $\mathbf{det}(\mathbf{adj}\,(A)) = (\mathbf{det}\,A)^{n-1}$.

Problem 6.4.13 Let $(A, B, S) \in (\mathbf{GL}_n(\mathbb{F}))^3$. Prove that

❶ $\mathbf{adj}\,(\mathbf{adj}\,(A)) = (\mathbf{det}\,A)^{n-2}A$.

❷ $\mathbf{adj}\,(AB) = \mathbf{adj}\,(A)\,\mathbf{adj}\,(B)$.

❸ $\mathbf{adj}\,(SAS^{-1}) = S(\mathbf{adj}\,(A))S^{-1}$.

Problem 6.4.14 If $A \in \mathbf{GL}_2(\mathbb{F})$, , and let k be a positive integer. Prove that $\mathbf{det}(\underbrace{\mathbf{adj}\cdots\mathbf{adj}}_{k}(A)) = \mathbf{det}\,A$.

Problem 6.4.15 Find the determinant

$$\det\begin{bmatrix} (b+c)^2 & ab & ac \\ ab & (a+c)^2 & bc \\ ac & bc & (a+b)^2 \end{bmatrix}$$

by hand, making explicit all your calculations.

Problem 6.4.16 The matrix

$$\begin{bmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{bmatrix}$$

is known as a *circulant matrix*. Prove that its determinant is $(a+b+c+d)(a-b+c-d)((a-c)^2+(b-d)^2)$.

6.5 Determinants and Linear Systems

313 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. The following are all equivalent

❶ $\mathbf{det}\,A \neq 0_{\mathbb{F}}$.

❷ A is invertible.

❸ There exists a unique solution $X \in \mathbf{M}_{n\times 1}(\mathbb{F})$ to the equation $AX = Y$.

❹ If $AX = \mathbf{0}_{n\times 1}$ then $X = \mathbf{0}_{n\times 1}$.

Proof: *We prove the implications in sequence:*
❶ $\implies$ ❷: *follows from Corollary 312*
❷ $\implies$ ❸: *If* A *is invertible and* $AX = Y$ *then* $X = A^{-1}Y$ *is the unique solution of this equation.*
❸ $\implies$ ❹: *follows by putting* $Y = \mathbf{0}_{n\times 1}$
❹ $\implies$ ❶: *Let* R *be the row echelon form of* A. *Since* $RX = \mathbf{0}_{n\times 1}$ *has only* $X = \mathbf{0}_{n\times 1}$ *as a solution, every entry on the diagonal of* R *must be non-zero,* R *must be triangular, and hence* $\mathbf{det}\, R \neq 0_{\mathbb{F}}$. *Since* $A = PR$ *where* P *is an invertible* $n \times n$ *matrix, we deduce that* $\mathbf{det}\, A = \mathbf{det}\, P\, \mathbf{det}\, R \neq 0_{\mathbb{F}}$.
❑

The contrapositive form of the implications ❶ and ❹ will be used later. Here it is for future reference.

314 Corollary Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. If there is $X \neq \mathbf{0}_{n\times 1}$ such that $AX = \mathbf{0}_{n\times 1}$ then $\mathbf{det}\, A = 0_{\mathbb{F}}$.

Homework

Problem 6.5.1 For which a is the matrix $\begin{bmatrix} -1 & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{bmatrix}$ singular (non-invertible)?

Chapter 7

Eigenvalues and Eigenvectors

7.1 Similar Matrices

315 Definition We say that $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ is *similar* to $B \in \mathbf{M}_{n\times n}(\mathbb{F})$ if there exist a matrix $P \in \mathbf{GL}_n(\mathbb{F})$ such that

$$B = PAP^{-1}.$$

316 Theorem Similarity is an equivalence relation.

> **Proof:** *Let* $A \in \mathbf{M}_{n\times n}(\mathbb{F})$. *Then* $A = \mathbf{I}_n A \mathbf{I}_n^{-1}$, *so similarity is reflexive. If* $B = PAP^{-1}$ *(*$P \in \mathbf{GL}_n(\mathbb{F})$ *) then* $A = P^{-1}BP$ *so similarity is symmetric. Finally, if* $B = PAP^{-1}$ *and* $C = QBQ^{-1}$ *(*$P \in \mathbf{GL}_n(\mathbb{F})$, $Q \in \mathbf{GL}_n(\mathbb{F})$*) then* $C = QPAP^{-1}Q^{-1} = QPA(QP)^{-1}$ *and so similarity is transitive.*
> ❑

Since similarity is an equivalence relation, it partitions the set of $n \times n$ matrices into equivalence classes by Theorem 29.

317 Definition A matrix is said to be *diagonalisable* if it is similar to a diagonal matrix.

Suppose that

$$A = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Then if K is a positive integer

$$A^K = \begin{bmatrix} \lambda_1^K & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^K & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n^K \end{bmatrix}.$$

In particular, if B is similar to A then

$$B^K = \underbrace{(PAP^{-1})(PAP^{-1})\cdots(PAP^{-1})}_{K \text{ factors}} = PA^KP^{-1} = P\begin{bmatrix} \lambda_1^K & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2^K & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n^K \end{bmatrix} P^{-1},$$

so we have a simpler way of computing B^K. Our task will now be to establish when a particular square matrix is diagonalisable.

7.2 Eigenvalues and Eigenvectors

Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ be a square diagonalisable matrix. Then there exist $P \in \mathbf{GL}_n(\mathbb{F})$ and a diagonal matrix $D \in \mathbf{M}_{n\times n}(\mathbb{F})$ such that $P^{-1}AP = D$, whence $AP = DP$. Put

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}, \quad P = [P_1; P_2; \cdots; P_n],$$

where the P_k are the columns of P. Then

$$AP = DP \implies [AP_1; AP_2; \cdots; AP_n] = [\lambda_1 P_1; \lambda_2 P_2; \cdots; \lambda_n P_n],$$

from where it follows that $AP_k = \lambda_k P_k$. This motivates the following definition.

318 Definition Let V be a finite-dimensional vector space over a field $\mathbb{F}$ and let $T : V \to V$ be a linear transformation. A scalar $\lambda \in \mathbb{F}$ is called an *eigenvalue* of T if there is a $\overrightarrow{v} \neq \overrightarrow{0}$ (called an *eigenvector*) such that $T(\overrightarrow{v}) = \lambda\overrightarrow{v}$.

319 Example Shew that if λ is an eigenvalue of $T : V \to V$, then λ^k is an eigenvalue of $T^k : V \to V$, for $k \in \mathbb{N} \setminus \{0\}$.

> **Solution:** ► *Assume that* $T(\overrightarrow{v}) = \lambda\overrightarrow{v}$. *Then*
>
> $$T^2(\overrightarrow{v}) = TT(\overrightarrow{v}) = T(\lambda\overrightarrow{v}) = \lambda T(\overrightarrow{v}) = \lambda(\lambda\overrightarrow{v}) = \lambda^2\overrightarrow{v}.$$
>
> *Continuing the iterations we obtain* $T^k(\overrightarrow{v}) = \lambda^k\overrightarrow{v}$, *which is what we want.* ◄

320 Theorem Let $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ be the matrix representation of $T : V \to V$. Then $\lambda \in \mathbb{F}$ is an eigenvalue of T if an only if $\mathbf{det}(\lambda\mathbf{I}_n - A) = 0_{\mathbb{F}}$.

> **Proof:** λ *is an eigenvalue of* A $\iff$ *there is* $\overrightarrow{v} \neq \overrightarrow{0}$ *such that* $A\overrightarrow{v} = \lambda\overrightarrow{v}$ $\iff$ $\lambda\overrightarrow{v} - A\overrightarrow{v} = \overrightarrow{0}$ $\iff$ $\lambda\mathbf{I}_n\overrightarrow{v} - A\overrightarrow{v} = \overrightarrow{0}$ $\iff$ $\mathbf{det}(\lambda\mathbf{I}_n - A) = 0_{\mathbb{F}}$ *by Corollary 314.*❑

321 Definition The equation

$$\mathbf{det}(\lambda\mathbf{I}_n - A) = 0_{\mathbb{F}}$$

is called the *characteristic equation of* A or *secular equation of* A. The polynomial $p(\lambda) = \mathbf{det}(\lambda\mathbf{I}_n - A)$ is the characteristic polynomial of A.

322 Example Let $A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$. Find

❶ The characteristic polynomial of A.

❷ The eigenvalues of A.

❸ The corresponding eigenvectors.

Solution: ▶ *We have*

❶

$$\begin{aligned} \mathbf{det}(\lambda \mathbf{I}_4 - A) &= \mathbf{det} \begin{bmatrix} \lambda - 1 & -1 & 0 & 0 \\ -1 & \lambda - 1 & 0 & 0 \\ 0 & 0 & \lambda - 1 & -1 \\ 0 & 0 & -1 & \lambda - 1 \end{bmatrix} \\ &= (\lambda - 1)\,\mathbf{det} \begin{bmatrix} \lambda - 1 & 0 & 0 \\ 0 & \lambda - 1 & -1 \\ 0 & -1 & \lambda - 1 \end{bmatrix} + \mathbf{det} \begin{bmatrix} -1 & 0 & 0 \\ 0 & \lambda - 1 & -1 \\ 0 & -1 & \lambda - 1 \end{bmatrix} \\ &= (\lambda - 1)((\lambda - 1)((\lambda - 1)^2 - 1)) + (-((\lambda - 1)^2 - 1)) \\ &= (\lambda - 1)((\lambda - 1)(\lambda - 2)(\lambda)) - (\lambda - 2)(\lambda) \\ &= (\lambda - 2)(\lambda)((\lambda - 1)^2 - 1) \\ &= (\lambda - 2)^2(\lambda)^2 \end{aligned}$$

❷ *The eigenvalues are clearly* $\lambda = 0$ *and* $\lambda = 2$.

❸ *If* $\lambda = 0$, *then*

$$0\mathbf{I}_4 - A = \begin{bmatrix} -1 & -1 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & -1 & -1 \end{bmatrix}.$$

This matrix has row-echelon form

$$\begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and if

$$\begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

then $\mathbf{c} = -\mathbf{d}$ *and* $\mathbf{a} = -\mathbf{b}$
Thus the general solution of the system $(0\mathbf{I}_4 - A)X = \mathbf{0}_{n\times 1}$ *is*

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \mathbf{a} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + \mathbf{c} \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}.$$

If $\lambda = 2$, *then*

$$\mathbf{2I}_4 - \mathbf{A} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

This matrix has row-echelon form

$$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and if

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

then $c = d$ *and* $a = b$
Thus the general solution of the system $(2\mathbf{I}_4 - A)X = \mathbf{0}_{n\times 1}$ *is*

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

Thus for $\lambda = 0$ *we have the eigenvectors*

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

and for $\lambda = 2$ *we have the eigenvectors*

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}.$$

◀

323 Theorem If $\lambda = 0_{\mathbb{F}}$ is an eigenvalue of A, then A is non-invertible.

Proof: *Put* $p(\lambda) = \mathbf{det}(\lambda\mathbf{I}_n - A)$. *Then* $p(0_{\mathbb{F}}) = \mathbf{det}(-A) = (-1)^n \mathbf{det}\, A$ *is the constant term of the characteristic polynomial. If* $\lambda = 0_{\mathbb{F}}$ *is an eigenvalue then*

$$p(0_{\mathbb{F}}) = 0_{\mathbb{F}} \implies \mathbf{det}\, A = 0_{\mathbb{F}},$$

and hence A *is non-invertible by Theorem 313.* ❑

324 Theorem Similar matrices have the same characteristic polynomial.

Proof: *We have*

$$\begin{aligned}\mathbf{det}(\lambda\mathbf{I}_n - SAS^{-1}) &= \mathbf{det}(\lambda S\mathbf{I}_n S^{-1} - SAS^{-1}) \\ &= \mathbf{det}\, S(\lambda\mathbf{I}_n - A)S^{-1} \\ &= (\mathbf{det}\, S)(\mathbf{det}(\lambda\mathbf{I}_n - A))(\mathbf{det}\, S^{-1}) \\ &= (\mathbf{det}\, S)(\mathbf{det}(\lambda\mathbf{I}_n - A))\left(\frac{1}{\mathbf{det}\, S}\right) \\ &= \mathbf{det}(\lambda\mathbf{I}_n - A),\end{aligned}$$

from where the result follows.❑

Homework

Problem 7.2.1 Find the eigenvalues and eigenvectors of

$$A = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Problem 7.2.2 Let A be a 2×2 matrix over some some field $\mathbb{F}$. Prove that the characteristic polynomial of A is

$$\lambda^2 - (\mathbf{tr}\,(A))\lambda + \mathbf{det}\, A.$$

Problem 7.2.3 A matrix $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$ satisfies $\mathbf{tr}\,(A) = -1$ and $\mathbf{det}\, A = -6$. Find the value of $\mathbf{det}(\mathbf{I}_2 + A)$.

Problem 7.2.4 A 2×2 matrix A with real entries has characteristic polynomial $p(\lambda) = \lambda^2 + 2\lambda - 1$. Find the value of $\mathbf{det}(2\mathbf{I}_2 + A)$.

Problem 7.2.5 Let $A = \begin{bmatrix} 0 & 2 & -1 \\ 2 & 3 & -2 \\ -1 & -2 & 0 \end{bmatrix}$. Find

❶ The characteristic polynomial of A.

❷ The eigenvalues of A.

❸ The corresponding eigenvectors.

Problem 7.2.6 Describe all matrices $A \in \mathbf{M}_{2\times 2}(\mathbb{R})$ having eigenvalues 1 and -1.

Problem 7.2.7 Let $A \in \mathbf{M}_{n\times n}(\mathbb{R})$. Demonstrate that A has the same characteristic polynomial as its transpose.

7.3 Diagonalisability

In this section we find conditions for diagonalisability.

325 Theorem Let $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\} \subset V$ be the eigenvectors corresponding to the *different* eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ (in that order). Then these eigenvectors are linearly independent.

Proof: *Let* $T : V \to V$ *be the underlying linear transformation. We proceed by induction. For* $k = 1$ *the result is clear. Assume that every set of* $k - 1$ *eigenvectors that correspond to* $k - 1$ *distinct eigenvalues is linearly independent and let the eigenvalues* $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ *have corresponding eigenvectors* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{k-1}$*. Let* λ *be a eigenvalue different from the* $\lambda_1, \lambda_2, \ldots, \lambda_{k-1}$ *and let its corresponding eigenvector be* $\vec{v}$*. If* $\vec{v}$ *were linearly dependent of the* $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_{k-1}$*, we would have*

$$x\vec{v} + x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_{k-1}\vec{v}_{k-1} = \vec{0}. \tag{7.1}$$

Now

$$T(x\vec{v} + x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_{k-1}\vec{v}_{k-1}) = T(\vec{0}) = \vec{0},$$

by Theorem 240. This implies that

$$x\lambda\vec{v} + x_1\lambda_1\vec{v}_1 + x_2\lambda_2\vec{v}_2 + \cdots + x_{k-1}\lambda_{k-1}\vec{v}_{k-1} = \vec{0}. \tag{7.2}$$

From 7.2 take away λ times 7.1, obtaining

$$x_1(\lambda_1 - \lambda)\overrightarrow{v}_1 + x_2(\lambda_2\overrightarrow{v}_2 + \cdots + x_{k-1}(\lambda_{k-1} - \lambda)\overrightarrow{v}_{k-1} = \overrightarrow{0} \tag{7.3}$$

Since $\lambda - \lambda_i \neq 0_{\mathbb{F}}$ 7.3 is saying that the eigenvectors $\overrightarrow{v}_1, \overrightarrow{v}_2, \ldots, \overrightarrow{v}_{k-1}$ are linearly dependent, a contradiction. Thus the claim follows for k distinct eigenvalues and the result is proven by induction. ❑

326 Theorem A matrix $A \in \mathbf{M}_{n\times n}(\mathbb{F})$ is diagonalisable if and only if it possesses n linearly independent eigenvectors.

Proof: *Assume first that A is diagonalisable, so there exists $P \in \mathbf{GL}_n(\mathbb{F})$ and*

$$D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

such that

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Then

$$[AP_1; AP_2; \cdots; AP_n] = AP = P\begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1 P_1; \lambda_2 P_2; \cdots; \lambda_n P_n],$$

where the P_k are the columns of P. Since P is invertible, the P_k are linearly independent by virtue of Theorems 204 and 313.

Conversely, suppose now that $\overrightarrow{v}_1, \ldots, \overrightarrow{v}_n$ are n linearly independent eigenvectors, with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Put

$$P = [\overrightarrow{v}_1; \ldots; \overrightarrow{v}_n], \quad D = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}.$$

Since $A\vec{v}_i = \lambda_i\vec{v}_i$ we see that $AP = PD$. Again P is invertible by Theorems 204 and 313 since the $\vec{v}_k$ are linearly independent. Left multiplying by P^{-1} we deduce $P^{-1}AP = D$, from where A is diagonalisable. ❑

327 Example Shew that the following matrix is diagonalisable:

$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

and find a diagonal matrix D and an invertible matrix P such that

$$A = PDP^{-1}.$$

Solution: ▶ *Verify that the characteristic polynomial of A is*

$$\lambda^3 - 3\lambda^2 - 4\lambda + 12 = (\lambda - 2)(\lambda + 2)(\lambda - 3).$$

The eigenvector for $\lambda = -2$ is

$$\begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}.$$

The eigenvector for $\lambda = 2$ is

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

The eigenvector for $\lambda = 3$ is

$$\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

We may take

$$D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & -1 & 4 \\ -1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix}.$$

We also find

$$P^{-1} = \begin{bmatrix} \frac{1}{5} & -1 & \frac{1}{5} \\ 0 & -1 & 1 \\ \frac{1}{5} & 0 & \frac{1}{5} \end{bmatrix}.$$

◀

Homework

Problem 7.3.1 Let A be a 2×2 matrix with eigenvalues 1 and -2 and corresponding eigenvectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, respectively. Determine A^{10}.

Problem 7.3.2 Consider the matrix $A = \begin{bmatrix} 9 & -4 \\ 20 & -9 \end{bmatrix}$.

1. Find the characteristic polynomial of A.
2. Find the eigenvalues of A.
3. Find the eigenvectors of A.
4. If $A^{20} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, find $a + d$.

Problem 7.3.3 Let $A \in \mathbf{M}_{3\times 3}(\mathbb{R})$ have characteristic polynomial

$$(\lambda + 1)^2(\lambda - 3).$$

One of the eigenvalues has two eigenvectors $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. The other eigenvalue has corresponding eigenvector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Determine A.

Problem 7.3.4 Let

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

1. Find $\mathbf{det}\, A$.
2. Find A^{-1}.
3. Find $\mathbf{rank}\,(A - \mathbf{I}_4)$.
4. Find $\mathbf{det}(A - \mathbf{I}_4)$.
5. Find the characteristic polynomial of A.
6. Find the eigenvalues of A.
7. Find the eigenvectors of A.
8. Find A^{10}.

Problem 7.3.5 Consider the matrix

$$A = \begin{bmatrix} 1 & a & 1 \\ 0 & 1 & b \\ 0 & 0 & c \end{bmatrix}.$$

❶ Find the characteristic polynomial of A.

❷ Explain whether A is diagonalisable when $a = 0, c = 1$.

❸ Explain whether A is diagonalisable when $a \neq 0, c = 1$.

❹ Explain whether A is diagonalisable when $c \neq 1$.

Problem 7.3.6 Find a closed formula for A^n, if

$$A = \begin{bmatrix} -7 & -6 \\ 12 & 10 \end{bmatrix}.$$

Problem 7.3.7 Let $U \in \mathbf{M}_{n\times n}(\mathbb{R})$ be a square matrix all whose entries are equal to 1.

1. Demonstrate that $U^2 = nU$.
2. Find $\mathbf{det}\, U$.
3. Prove that $\mathbf{det}(\lambda \mathbf{I}_n - U) = \lambda^{n-1}(\lambda - n)$.
4. Shew that $\mathbf{dim}\,\mathbf{ker}\,(U) = n - 1$.
5. Shew that

$$U = P \begin{bmatrix} n & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} P^{-1},$$

where

$$P = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & -1 & -1 & \cdots & -1 & -1 \end{bmatrix}.$$

7.4 Theorem of Cayley and Hamilton

328 Theorem (Cayley-Hamilton) A matrix $A \in \mathbf{M}_n(\mathbb{F})$ satisfies its characteristic polynomial.

Proof: *Put* $B = \lambda\mathbf{I}_n - A$. *We can write*

$$\mathbf{det}\, B = \mathbf{det}(\lambda\mathbf{I}_n - A) = \lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_n,$$

as $\mathbf{det}(\lambda\mathbf{I}_n - A)$ *is a polynomial of degree* n.

Since $\mathbf{adj}\,(B)$ *is a matrix obtained by using* $(n-1)\times(n-1)$ *determinants from* B, *we may write*

$$\mathbf{adj}\,(B) = \lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + B_0.$$

Hence

$$\mathbf{det}(\lambda\mathbf{I}_n - A)\mathbf{I}_n = (B)(\mathbf{adj}\,(B)) = (\lambda\mathbf{I}_n - A)(\mathbf{adj}\,(B)),$$

from where

$$\lambda^n\mathbf{I}_n + b_1\mathbf{I}_n\lambda^{n-1} + b_2\mathbf{I}_n\lambda^{n-2} + \cdots + b_n\mathbf{I}_n = (\lambda\mathbf{I}_n - A)(\lambda^{n-1}B_{n-1} + \lambda^{n-2}B_{n-2} + \cdots + B_0).$$

By equating coefficients we deduce

$$\begin{aligned} \mathbf{I}_n &= B_{n-1} \\ b_1\mathbf{I}_n &= -AB_{n-1} + B_{n-2} \\ b_2\mathbf{I}_n &= -AB_{n-2} + B_{n-3} \\ &\vdots \\ b_{n-1}\mathbf{I}_n &= -AB_1 + B_0 \\ b_n\mathbf{I}_n &= -AB_0. \end{aligned}$$

Multiply now the k-th row by A^{n-k} (the first row appearing is really the 0-th row):

$$\begin{aligned} A^n &= A^n B_{n-1} \\ b_1 A^{n-1} &= -A^n B_{n-1} + A^{n-1} B_{n-2} \\ b_2 A^{n-2} &= -A^{n-1} B_{n-2} + A^{n-2} B_{n-3} \\ &\vdots \\ b_{n-1} A &= -A^2 B_1 + AB_0 \\ b_n \mathbf{I}_n &= -AB_0. \end{aligned}$$

Add all the rows and through telescopic cancellation obtain

$$A^n + b_1 A^{n-1} + \cdots + b_{n-1} A + b_n \mathbf{I}_n = \mathbf{0}_n,$$

from where the theorem follows. ❑

329 Example From example 327 the matrix

$$\begin{bmatrix} 1 & -1 & -1 \\ 1 & 3 & 1 \\ -3 & 1 & -1 \end{bmatrix}$$

has characteristic polynomial

$$(\lambda - 3)(\lambda - 2)(\lambda + 2) = \lambda^3 - 3\lambda^2 - 4\lambda + 12,$$

hence the inverse of this matrix can be obtained by observing that

$$A^3 - 3A^2 - 4A + 12\mathbf{I}_3 = \mathbf{0}_3 \implies A^{-1} = -\frac{1}{12}(A^2 - 3A - 4\mathbf{I}_3) = \begin{bmatrix} 1/3 & 1/6 & -1/6 \\ 1/6 & 1/3 & 1/6 \\ -5/6 & -1/6 & -1/3 \end{bmatrix}.$$

Homework

Problem 7.4.1 A 3×3 matrix A has characteristic polynomial $\lambda(\lambda - 1)(\lambda + 2)$. What is the characteristic polynomial of A^2?

Chapter 8

Linear Algebra and Geometry

8.1 Points and Bi-points in $\mathbb{R}^2$

$\mathbb{R}^2$ is the set of all points $\mathbf{A} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$ with real number coordinates on the plane, as in figure 8.1. We use the notation $\mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to denote the *origin*.

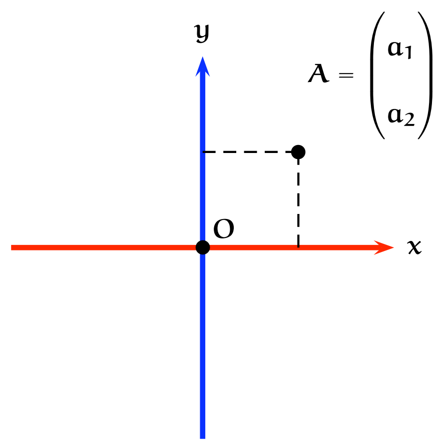


Figure 8.1: Rectangular coordinates in $\mathbb{R}^2$.

Given $\mathbf{A} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \in \mathbb{R}^2$ and $\mathbf{B} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \in \mathbb{R}^2$ we define their addition as

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}. \tag{8.1}$$

Similarly, we define the scalar multiplication of a point of $\mathbb{R}^2$ by the scalar $\alpha \in \mathbb{R}$ as

$$\alpha A = \alpha \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \end{pmatrix}. \tag{8.2}$$

☞ *Throughout this chapter, unless otherwise noted, we will use the convention that a point* $\mathbf{A} \in \mathbb{R}^2$ *will have its coordinates named after its letter, thus*

$$\mathbf{A} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}.$$

330 Definition Consider the points $\mathbf{A} \in \mathbb{R}^2, \mathbf{B} \in \mathbb{R}^2$. By the *bi-point* starting at $\mathbf{A}$ and ending at $\mathbf{B}$, denoted by $[\mathbf{A}, \mathbf{B}]$, we mean the directed line segment from $\mathbf{A}$ to $\mathbf{B}$. We define

$$[\mathbf{A}, \mathbf{A}] = \mathbf{O} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

☞ *The bi-point* $[\mathbf{A}, \mathbf{B}]$ *can be thus interpreted as an arrow starting at* $\mathbf{A}$ *and finishing, with the arrow tip, at* $\mathbf{B}$*. We say that* $\mathbf{A}$ *is the* tail *of the bi-point* $[\mathbf{A}, \mathbf{B}]$ *and that* $\mathbf{B}$ *is its* head*. Some authors use the terminology* "fixed vector" *instead of "bi-point."*

331 Definition Let $\mathbf{A} \neq \mathbf{B}$ be points on the plane and let $\mathbf{L}$ be the line passing through $\mathbf{A}$ and $\mathbf{B}$. The *direction* of the bi-point $[\mathbf{A}, \mathbf{B}]$ is the direction of the line $\mathbf{L}$, that is, the angle $\theta \in \left]-\frac{\pi}{2}; \frac{\pi}{2}\right]$ that the line $\mathbf{L}$ makes with the horizontal. See figure 8.2.

332 Definition Let $\mathbf{A}, \mathbf{B}$ lie on line $\mathbf{L}$, and let $\mathbf{C}, \mathbf{D}$ lie on line $\mathbf{L}'$. If $\mathbf{L} || \mathbf{L}'$ then we say that $[\mathbf{A}, \mathbf{B}]$ has the same direction as $[\mathbf{C}, \mathbf{D}]$. We say that the bi-points $[\mathbf{A}, \mathbf{B}]$ and $[\mathbf{C}, \mathbf{D}]$ have the *same sense* if they have the same direction and if both their heads lie on the same half-plane made by the line joining their tails. They have *opposite sense* if they have the same direction and if both their heads lie on alternative half-planes made by the line joining their tails. See figures 8.3 and 8.4 .

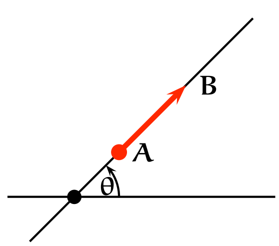


Figure 8.2: Direction of a bi-point

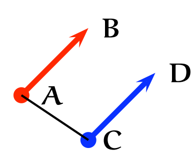


Figure 8.3: Bi-points with the same sense.

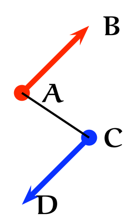


Figure 8.4: Bi-points with opposite sense.

☞ *Bi-point* $[B, A]$ *has the opposite sense of* $[A, B]$ *and so we write*

$$[B, A] = -[A, B].$$

333 Definition Let $A \neq B$. The *Euclidean length or norm* of bi-point $[A, B]$ is simply the distance between A and B and it is denoted by

$$||[A, B]|| = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}.$$

We define

$$||[A, A]|| = ||O|| = 0.$$

A bi-point is said to have *unit length* if it has norm 1.

☞ *A bi-point is completely determined by three things: (i) its norm, (ii) its direction, and (iii) its sense.*

334 Definition (Chasles' Rule) Two bi-points are said to be *contiguous* if one has as tail the head of the other. In such case we define the sum of contiguous bi-points $[A, B]$ and $[B, C]$ by *Chasles' Rule*

$$[A, B] + [B, C] = [A, C].$$

See figure 8.5.

335 Definition (Scalar Multiplication of Bi-points) Let $\lambda \in \mathbb{R} \setminus \{0\}$ and $A \neq B$. We define

$$0[A, B] = O$$

and

$$\lambda[A, A] = O.$$

We define $\lambda[A, B]$ as follows.

1. $\lambda[A, B]$ has the direction of $[A, B]$.
2. $\lambda[A, B]$ has the sense of $[A, B]$ if $\lambda > 0$ and sense opposite $[A, B]$ if $\lambda < 0$.
3. $\lambda[A, B]$ has norm $|\lambda|||[A, B]||$ which is a contraction of $[A, B]$ if $0 < |\lambda| < 1$ or a dilatation of $[A, B]$ if $|\lambda| > 1$.

See figure 8.6 for some examples.

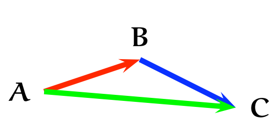


Figure 8.5: Chasles' Rule.

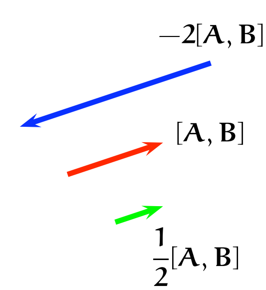


Figure 8.6: Scalar multiplication of bi-points.

8.2 Vectors in $\mathbb{R}^2$

336 Definition (Midpoint) Let A, B be points in $\mathbb{R}^2$. We define the *midpoint* of the bi-point $[A, B]$ as

$$\frac{A+B}{2} = \begin{pmatrix} \frac{a_1+b_1}{2} \\ \frac{a_2+b_2}{2} \end{pmatrix}.$$

337 Definition (Equipollence) Two bi-points $[X, Y]$ and $[A, B]$ are said to be *equipollent* written $[X, Y] \sim [A, B]$ if the midpoints of the bi-points $[X, B]$ and $[Y, A]$ coincide, that is,

$$[X, Y] \sim [A, B] \Leftrightarrow \frac{X+B}{2} = \frac{Y+A}{2}.$$

See figure 8.7.

Geometrically, equipollence means that the quadrilateral $XYBA$ is a parallelogram. Thus the bi-points $[X, Y]$ and $[A, B]$ have the same norm, sense, and direction.

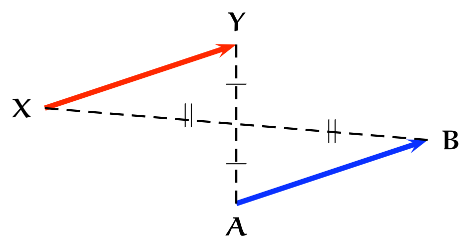


Figure 8.7: Equipollent bi-points.

338 Lemma Two bi-points $[X, Y]$ and $[A, B]$ are equipollent if and only if

$$\begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix}.$$

Proof: *This is immediate, since*

$$\begin{aligned} [X, Y] \sim [A, B] \quad & \iff \quad \begin{pmatrix} \frac{a_1+y_1}{2} \\ \frac{a_2+y_2}{2} \end{pmatrix} = \begin{pmatrix} \frac{b_1+x_1}{2} \\ \frac{b_2+x_2}{2} \end{pmatrix} \\ & \iff \quad \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix}, \end{aligned}$$

as desired. ❑

☞ *From Lemma 338, equipollent bi-points have the same norm, the same direction, and the same sense.*

339 Theorem Equipollence is an equivalence relation.

Proof: *Write* $[X, Y] \sim [A, B]$ *if* $[X, Y]$ *if equipollent to* $[A, B]$. *Now* $[X, Y] \sim [X, Y]$ *since* $\begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} = \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix}$ *and so the relation is reflexive. Also*

$$\begin{aligned} [X, Y] \sim [A, B] &\iff \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} \\ &\iff \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} = \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} \\ &\iff [A, B] \sim [X, Y], \end{aligned}$$

and the relation is symmetric. Finally

$$\begin{aligned} [X, Y] \sim [A, B] \wedge [A, B] \sim [U, V] &\iff \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} = \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} \\ &\qquad \wedge \begin{pmatrix} b_1 - a_1 \\ b_2 - a_2 \end{pmatrix} = \begin{pmatrix} v_1 - u_1 \\ v_2 - au_2 \end{pmatrix} \\ &\iff \begin{pmatrix} y_1 - x_1 \\ y_2 - x_2 \end{pmatrix} = \begin{pmatrix} v_1 - u_1 \\ v_2 - u_2 \end{pmatrix} \\ &\iff [X, Y] \sim [U, V], \end{aligned}$$

and the relation is transitive. ❑

340 Definition (Vectors on the Plane) The equivalence class in which the bi-point $[X, Y]$ falls is called the *vector* (or *free vector*) from X to Y, and is denoted by $\overrightarrow{XY}$. Thus we write

$$[X, Y] \in \overrightarrow{XY} = \begin{bmatrix} y_1 - x_1 \\ y_2 - x_2 \end{bmatrix}.$$

If we desire to talk about a vector without mentioning a bi-point representative, we write, say, $\overrightarrow{v}$, thus denoting vectors with boldface lowercase letters. If it is necessary to mention the coordinates of $\overrightarrow{v}$ we will write

$$\overrightarrow{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}.$$

☞ *For any point* X *on the plane, we have* $\overrightarrow{XX} = \overrightarrow{0}$, *the* zero vector. *If* $[X, Y] \in \overrightarrow{v}$ *then* $[Y, X] \in -\overrightarrow{v}$.

341 Definition (Position Vector) For any particular point $P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \in \mathbb{R}^2$ we may form the vector $\overrightarrow{OP} = \begin{bmatrix} p_1 \\ p_2 \end{bmatrix}$. We call $\overrightarrow{OP}$ the *position vector* of P and we use boldface lowercase letters to denote the equality $\overrightarrow{OP} = \overrightarrow{p}$.

342 Example The vector into which the bi-point with tail at $A = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and head at $B = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ falls is

$$\overrightarrow{AB} = \begin{bmatrix} 3 - (-1) \\ 4 - 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}.$$

343 Example The bi-points $[A, B]$ and $[X, Y]$ with

$$A = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, B = \begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

$$X = \begin{pmatrix} 3 \\ 7 \end{pmatrix}, Y = \begin{pmatrix} 7 \\ 9 \end{pmatrix}$$

represent the same vector

$$\overrightarrow{AB} = \begin{bmatrix} 3 - (-1) \\ 4 - 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = \begin{bmatrix} 7 - 3 \\ 9 - 7 \end{bmatrix} = \overrightarrow{XY}.$$

In fact, if $S = \begin{pmatrix} -1 + n \\ 2 + m \end{pmatrix}, T = \begin{pmatrix} 3 + n \\ 4 + m \end{pmatrix}$ then the infinite number of bi-points $[S, T]$ are representatives of of the vectors $\overrightarrow{AB} = \overrightarrow{XY} = \overrightarrow{ST}$.

Given two vectors $\overrightarrow{u}, \overrightarrow{v}$ we define their sum $\overrightarrow{u} + \overrightarrow{v}$ as follows. Find a bi-point representative $\overrightarrow{AB} \in \overrightarrow{u}$ and a contiguous bi-point representative $\overrightarrow{BC} \in \overrightarrow{v}$. Then by Chasles' Rule

$$\overrightarrow{u} + \overrightarrow{v} = \overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}.$$

Again, by virtue of Chasles' Rule we then have

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = -\overrightarrow{OA} + \overrightarrow{OB} = \overrightarrow{b} - \overrightarrow{a} \tag{8.3}$$

Similarly we define scalar multiplication of a vector by scaling one of its bi-point representatives.We define the norm of a vector $\overrightarrow{v} \in \mathbb{R}^2$ to be the norm of any of its bi-point representatives.

Componentwise we may see that given vectors $\overrightarrow{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, $\overrightarrow{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$, and a scalar $\lambda \in \mathbb{R}$ then their sum and scalar multiplication take the form

$$\overrightarrow{u} + \overrightarrow{v} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad \lambda\overrightarrow{u} = \begin{bmatrix} \lambda u_1 \\ \lambda u_2 \end{bmatrix}.$$

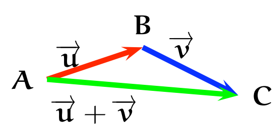


Figure 8.8: Addition of Vectors.

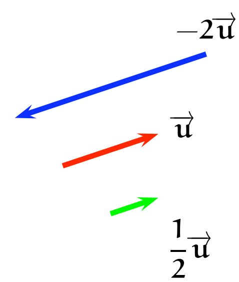


Figure 8.9: Scalar multiplication of vectors.

344 Example Diagonals are drawn in a rectangle $ABCD$. If $\overrightarrow{AB} = \overrightarrow{x}$ and $\overrightarrow{AC} = \overrightarrow{y}$, then $\overrightarrow{BC} = \overrightarrow{y} - \overrightarrow{x}$, $\overrightarrow{CD} = -\overrightarrow{x}$, $\overrightarrow{DA} = \overrightarrow{x} - \overrightarrow{y}$, and $\overrightarrow{BD} = \overrightarrow{y} - 2\overrightarrow{x}$.

345 Definition (Parallel Vectors) Two vectors $\overrightarrow{u}$ and $\overrightarrow{v}$ are said to be *parallel* if there is a scalar λ such that $\overrightarrow{u} = \lambda\overrightarrow{v}$. If $\overrightarrow{u}$ is parallel to $\overrightarrow{v}$ we write $\overrightarrow{u} || \overrightarrow{v}$. We denote by $\mathbb{R}\overrightarrow{v} = \{\alpha\overrightarrow{v} : \alpha \in \mathbb{R}\}$, the set of all vectors parallel to $\overrightarrow{v}$.

☞ $\overrightarrow{0}$ *is parallel to every vector.*

346 Definition If $\overrightarrow{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$, then we define its *norm* as $||\overrightarrow{u}|| = \sqrt{u_1^2 + u_2^2}$. The distance between two vectors $\overrightarrow{u}$ and $\overrightarrow{v}$ is $\mathbf{d}\langle\overrightarrow{u}, \overrightarrow{v}\rangle = ||\overrightarrow{u} - \overrightarrow{v}||$.

347 Example Let $a \in \mathbb{R}$, $a > 0$ and let $\overrightarrow{v} \neq \overrightarrow{0}$. Find a vector with norm a and parallel to $\overrightarrow{v}$.

Solution: ► *Observe that* $\dfrac{\overrightarrow{v}}{||\overrightarrow{v}||}$ *has norm* 1 *as*

$$\left|\left|\frac{\overrightarrow{v}}{||\overrightarrow{v}||}\right|\right| = \frac{||\overrightarrow{v}||}{||\overrightarrow{v}||} = 1.$$

Hence the vector $a\dfrac{\overrightarrow{v}}{||\overrightarrow{v}||}$ *has norm* a *and it is in the direction of* $\overrightarrow{v}$. *One may also take* $-a\dfrac{\overrightarrow{v}}{||\overrightarrow{v}||}$.
◄

348 Example If M is the midpoint of the bi-point $[X, Y]$ then $\overrightarrow{XM} = \overrightarrow{MY}$ from where $\overrightarrow{XM} = \frac{1}{2}\overrightarrow{XY}$. Moreover, if T is any point, by Chasles' Rule

$$\begin{aligned} \overrightarrow{TX} + \overrightarrow{TY} &= \overrightarrow{TM} + \overrightarrow{MX} + \overrightarrow{TM} + \overrightarrow{MY} \\ &= 2\overrightarrow{TM} - \overrightarrow{XM} + \overrightarrow{MY} \\ &= 2\overrightarrow{TM}. \end{aligned}$$

349 Example Let $\triangle ABC$ be a triangle on the plane. Prove that the line joining the midpoints of two sides of the triangle is parallel to the third side and measures half its length.

Solution: ► *Let the midpoints of* $[A, B]$ *and* $[A, C]$ *be* M_C *and* M_B, *respectively. We shew that* $\overrightarrow{BC} = 2\overrightarrow{M_CM_B}$. *We have* $2\overrightarrow{AM_C} = \overrightarrow{AB}$ *and* $2\overrightarrow{AM_B} = \overrightarrow{AC}$. *Thus*

$$\begin{aligned} \overrightarrow{BC} &= \overrightarrow{BA} + \overrightarrow{AC} \\ &= -\overrightarrow{AB} + \overrightarrow{AC} \\ &= -2\overrightarrow{AM_C} + 2\overrightarrow{AM_B} \\ &= 2\overrightarrow{M_CA} + 2\overrightarrow{AM_B} \\ &= 2(\overrightarrow{M_CA} + \overrightarrow{AM_B}) \\ &= 2\overrightarrow{M_CM_B}, \end{aligned}$$

as we wanted to shew. ◄

350 Example In $\triangle ABC$, let M_C be the midpoint of side AB. Shew that

$$\overrightarrow{CM_C} = \frac{1}{2}\left(\overrightarrow{CA} + \overrightarrow{CB}\right).$$

Solution: ► *Since* $\overrightarrow{AM_C} = \overrightarrow{M_CB}$, *we have*

$$\begin{aligned} \overrightarrow{CA} + \overrightarrow{CB} &= \overrightarrow{CM_C} + \overrightarrow{M_CA} + \overrightarrow{CM_C} + \overrightarrow{M_CB} \\ &= 2\overrightarrow{CM_C} - \overrightarrow{AM_C} + \overrightarrow{M_CB} \\ &= 2\overrightarrow{CM_C}, \end{aligned}$$

which yields the desired result. ◄

351 Theorem (Section Formula) Let APB be a straight line and λ and μ be real numbers such that

$$\frac{||[A, P]||}{||[P, B]||} = \frac{\lambda}{\mu}.$$

With $\overrightarrow{a} = \overrightarrow{OA}$, $\overrightarrow{b} = \overrightarrow{OB}$, and $\overrightarrow{p} = \overrightarrow{OP}$, then

$$\overrightarrow{p} = \frac{\lambda\overrightarrow{b} + \mu\overrightarrow{a}}{\lambda + \mu}. \tag{8.4}$$

Proof: *Using Chasles' Rule for vectors,*

$$\overrightarrow{AB} = \overrightarrow{AO} + \overrightarrow{OB} = -\overrightarrow{a} + \overrightarrow{b},$$

$$\overrightarrow{AP} = \overrightarrow{AO} + \overrightarrow{OP} = -\overrightarrow{a} + \overrightarrow{p}.$$

Also, using Chasles' Rule for bi-points,

$$[A, P]\mu = \lambda([P, B]) = \lambda([P, A] + [A, B]) = \lambda(-[A, P] + [A, B]),$$

whence

$$[A, P] = \frac{\lambda}{\lambda + \mu}[A, B] \implies \overrightarrow{AP} = \frac{\lambda}{\lambda + \mu}\overrightarrow{AB} \implies \overrightarrow{p} - \overrightarrow{a} = \frac{\lambda}{\lambda + \mu}(\overrightarrow{b} - \overrightarrow{a}).$$

On combining these formulæ

$$(\lambda + \mu)(\overrightarrow{p} - \overrightarrow{a}) = \lambda(\overrightarrow{b} - \overrightarrow{a}) \implies (\lambda + \mu)\overrightarrow{p} = \lambda\overrightarrow{b} + \mu\overrightarrow{a},$$

from where the result follows. ❑

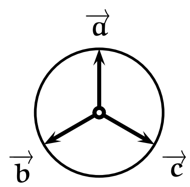


Figure 8.10: [A]. Problem 8.2.6.

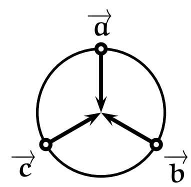


Figure 8.11: [B]. Problem 8.2.6.

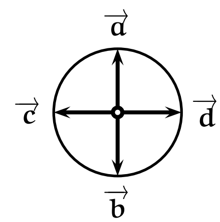


Figure 8.12: [C]. Problem 8.2.6.

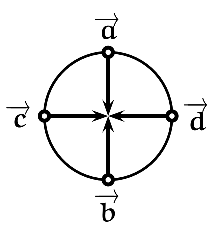


Figure 8.13: [D]. Problem 8.2.6.

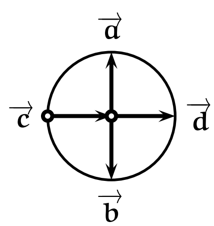


Figure 8.14: [E]. Problem 8.2.6.

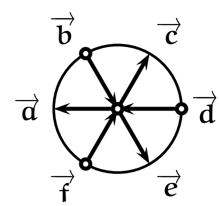


Figure 8.15: [F]. Problem 8.2.6.

Homework

Problem 8.2.1 Let a be a real number. Find the distance between $\begin{bmatrix} 1 \\ a \end{bmatrix}$ and $\begin{bmatrix} 1 - a \\ 1 \end{bmatrix}$.

Problem 8.2.2 Find all scalars λ for which $||\lambda\overrightarrow{v}|| = \frac{1}{2}$, where $\overrightarrow{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Problem 8.2.3 Given a pentagon $ABCDE$, find $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CD} + \overrightarrow{DE} + \overrightarrow{EA}$.

Problem 8.2.4 For which values of a will the vectors

$$\overrightarrow{a} = \begin{bmatrix} a+1 \\ a^2-1 \end{bmatrix}, \quad \overrightarrow{b} = \begin{bmatrix} 2a+5 \\ a^2-4a+3 \end{bmatrix}$$

will be parallel?

Problem 8.2.5 In $\triangle ABC$ let the midpoints of $[A,B]$ and $[A,C]$ be M_C and M_B, respectively. Put $\overrightarrow{M_C B} = \overrightarrow{x}$, $\overrightarrow{M_B C} = \overrightarrow{y}$, and $\overrightarrow{CA} = \overrightarrow{z}$. Express [A] $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{M_C M_B}$, [B] $\overrightarrow{AM_C} + \overrightarrow{M_C M_B} + \overrightarrow{M_B C}$, [C] $\overrightarrow{AC} + \overrightarrow{M_C A} - \overrightarrow{BM_B}$ in terms of $\overrightarrow{x}$, $\overrightarrow{y}$, and $\overrightarrow{z}$.

Problem 8.2.6 A circle is divided into three, four equal, or six equal parts (figures 8.10 through 8.15). Find the sum of the vectors. Assume that the divisions start or stop at the centre of the circle, as suggested in the figures.

Problem 8.2.7 Diagonals are drawn in a square (figures **??** through **??**). Find the vectorial sum $\overrightarrow{a} + \overrightarrow{b} + \overrightarrow{c}$. Assume that the diagonals either start, stop, or pass through the centre of the square, as suggested by the figures.

Problem 8.2.8 Prove that the mid-points of the sides of a skew quadrilateral form the vertices of a parallelogram.

Problem 8.2.9 $ABCD$ is a parallelogram. E is the mid-point of $[B,C]$ and F is the midpoint of $[D,C]$. Prove that

$$\overrightarrow{AC} + \overrightarrow{BD} = 2\overrightarrow{BC}.$$

Problem 8.2.10 Let A, B be two points on the plane. Construct two points I and J such that

$$\overrightarrow{IA} = -3\overrightarrow{IB}, \quad \overrightarrow{JA} = -\frac{1}{3}\overrightarrow{JB},$$

and then demonstrate that for any arbitrary point M on the plane

$$\overrightarrow{MA} + 3\overrightarrow{MB} = 4\overrightarrow{MI}$$

and

$$3\overrightarrow{MA} + \overrightarrow{MB} = 4\overrightarrow{MJ}.$$

Problem 8.2.11 You find an ancient treasure map in your great-grandfather's sea-chest. The sketch indicates that from the gallows you should walk to the oak tree, turn right $90°$ and walk a like distance, putting and x at the point where you stop; then go back to the gallows, walk to the pine tree, turn left $90°$, walk the same distance, mark point Y. Then you will find the treasure at the midpoint of the segment $\overline{XY}$. So you charter a sailing vessel and go to the remote south-seas island. On arrival, you readily locate the oak and pine trees, but unfortunately, the gallows was struck by lightning, burned to dust and dispersed to the winds. No trace of it remains. What do you do?

8.3 Dot Product in $\mathbb{R}^2$

352 Definition Let $(\overrightarrow{a}, \overrightarrow{b}) \in (\mathbb{R}^2)^2$. The *dot product* $\overrightarrow{a} \bullet \overrightarrow{b}$ of $\overrightarrow{a}$ and $\overrightarrow{b}$ is defined by

$$\overrightarrow{a} \bullet \overrightarrow{b} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \bullet \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = a_1 b_1 + a_2 b_2.$$

The following properties of the dot product are easy to deduce from the definition.

DP1 **Bilinearity**

$$(\overrightarrow{x} + \overrightarrow{y}) \bullet \overrightarrow{z} = \overrightarrow{x} \bullet \overrightarrow{z} + \overrightarrow{y} \bullet \overrightarrow{z}, \quad \overrightarrow{x} \bullet (\overrightarrow{y} + \overrightarrow{z}) = \overrightarrow{x} \bullet \overrightarrow{y} + \overrightarrow{x} \bullet \overrightarrow{z} \tag{8.5}$$

DP2 **Scalar Homogeneity**

$$(\alpha \overrightarrow{x}) \bullet \overrightarrow{y} = \overrightarrow{x} \bullet (\alpha \overrightarrow{y}) = \alpha(\overrightarrow{\mathbf{x}} \bullet \overrightarrow{\mathbf{y}}), \ \alpha \in \mathbb{R}. \tag{8.6}$$

DP3 **Commutativity**

$$\overrightarrow{x} \bullet \overrightarrow{y} = \overrightarrow{\mathbf{y}} \bullet \overrightarrow{\mathbf{x}} \tag{8.7}$$

DP4

$$\overrightarrow{x} \bullet \overrightarrow{x} \ \geq \ 0 \tag{8.8}$$

DP5

$$\overrightarrow{x} \bullet \overrightarrow{x} = 0 \Leftrightarrow \overrightarrow{x} = \overrightarrow{\mathbf{0}} \tag{8.9}$$

DP6

$$||\mathbf{x}|| = \sqrt{\overrightarrow{x}\bullet\overrightarrow{x}} \tag{8.10}$$

353 Example If we put

$$\overrightarrow{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

then we can write any vector $\overrightarrow{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ as a sum

$$\overrightarrow{a} = a_1\overrightarrow{i} + a_2\overrightarrow{j}.$$

The vectors

$$\overrightarrow{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \overrightarrow{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

satisfy $\overrightarrow{i}\bullet\overrightarrow{j} = 0$, and $\left|\left|\overrightarrow{i}\right|\right| = \left|\left|\overrightarrow{j}\right|\right| = 1$.

354 Definition Given vectors $\overrightarrow{a}$ and $\overrightarrow{b}$, we define the angle between them, denoted by $\widehat{(\overrightarrow{a}, \overrightarrow{b})}$, as the angle between any two contiguous bi-point representatives of $\overrightarrow{a}$ and $\overrightarrow{b}$.

355 Theorem

$$\overrightarrow{\mathbf{a}}\bullet\overrightarrow{b} = ||\overrightarrow{a}||||\overrightarrow{b}|| \cos \widehat{(\overrightarrow{a}, \overrightarrow{b})}.$$

Proof: *Using Al-Kashi's Law of Cosines on the length of the vectors, we have*

$$\begin{aligned}
&||\overrightarrow{b} - \overrightarrow{a}||^2 = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}|| \cos \widehat{(\overrightarrow{a}, \overrightarrow{b})} \\
&\Leftrightarrow (\overrightarrow{b} - \overrightarrow{a})\bullet(\overrightarrow{b} - \overrightarrow{a}) = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}|| \cos \widehat{(\overrightarrow{a}, \overrightarrow{b})} \\
&\Leftrightarrow \overrightarrow{b}\bullet\overrightarrow{b} - 2\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{a}\bullet\overrightarrow{a} = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}|| \cos \widehat{(\overrightarrow{a}, \overrightarrow{b})} \\
&\Leftrightarrow ||\overrightarrow{\mathbf{b}}||^2 - 2\overrightarrow{\mathbf{a}}\bullet\overrightarrow{\mathbf{b}} + ||\overrightarrow{\mathbf{b}}||^2 = ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2 - 2||\overrightarrow{a}||||\overrightarrow{b}|| \cos \widehat{(\overrightarrow{a}, \overrightarrow{b})} \\
&\Leftrightarrow \overrightarrow{a}\bullet\overrightarrow{b} = ||\overrightarrow{a}||||\overrightarrow{b}|| \cos \widehat{(\overrightarrow{a}, \overrightarrow{b})},
\end{aligned}$$

as we wanted to shew. ❑

Putting $\widehat{(\overrightarrow{a}, \overrightarrow{b})} = \frac{\pi}{2}$ in Theorem 355 we obtain the following corollary.

356 Corollary Two vectors in $\mathbb{R}^2$ are perpendicular if and only if their dot product is 0.

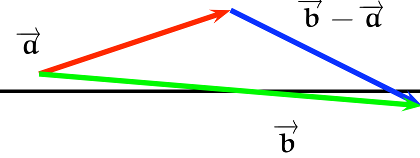


Figure 8.19: Theorem 355.

357 Definition Two vectors are said to be *orthogonal* if they are perpendicular. If $\overrightarrow{a}$ is orthogonal to $\overrightarrow{b}$, we write $\overrightarrow{a} \perp \overrightarrow{b}$.

358 Definition If $\overrightarrow{a} \perp \overrightarrow{b}$ and $||\overrightarrow{a}|| = \left|\left|\overrightarrow{b}\right|\right| = 1$ we say that $\overrightarrow{a}$ and $\overrightarrow{b}$ are *orthonormal*.

☞ *It follows that the vector $\overrightarrow{0}$ is simultaneously parallel and perpendicular to any vector!*

359 Definition Let $\overrightarrow{a} \in \mathbb{R}^2$ be fixed. Then the *orthogonal space* to $\overrightarrow{a}$ is defined and denoted by

$$\overrightarrow{a}^{\perp} = \{\overrightarrow{x} \in \mathbb{R}^2 : \overrightarrow{x} \perp \overrightarrow{a}\}.$$

Since $|\cos\theta| \leq 1$ we also have

360 Corollary (Cauchy-Bunyakovsky-Schwarz Inequality)

$$\left|\overrightarrow{a}\bullet\overrightarrow{b}\right| \leq ||\overrightarrow{a}||\left|\left|\overrightarrow{b}\right|\right|.$$

361 Corollary (Triangle Inequality)

$$\left|\left|\overrightarrow{a}+\overrightarrow{b}\right|\right| \leq ||\overrightarrow{a}|| + \left|\left|\overrightarrow{b}\right|\right|.$$

Proof:

$$\begin{aligned} ||\overrightarrow{a}+\overrightarrow{b}||^2 &= (\overrightarrow{a}+\overrightarrow{b})\bullet(\overrightarrow{a}+\overrightarrow{b}) \\ &= \overrightarrow{a}\bullet\overrightarrow{a} + 2\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{b}\bullet\overrightarrow{b} \\ &\leq ||\overrightarrow{a}||^2 + 2||\overrightarrow{a}||||\overrightarrow{b}|| + ||\overrightarrow{b}||^2 \\ &= (||\overrightarrow{a}|| + ||\overrightarrow{b}||)^2, \end{aligned}$$

from where the desired result follows. ❑

362 Corollary (Pythagorean Theorem) If $\overrightarrow{a} \perp \overrightarrow{b}$ then

$$||\overrightarrow{a}+\overrightarrow{b}||^2 = ||\overrightarrow{a}||^2 + \left|\left|\overrightarrow{b}\right|\right|^2.$$

Proof: *Since $\overrightarrow{a}\bullet\overrightarrow{b} = 0$, we have*

$$\begin{aligned} ||\overrightarrow{a}+\overrightarrow{b}||^2 &= (\overrightarrow{a}+\overrightarrow{b})\bullet(\overrightarrow{a}+\overrightarrow{b}) \\ &= \overrightarrow{a}\bullet\overrightarrow{a} + 2\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{b}\bullet\overrightarrow{b} \\ &= \overrightarrow{a}\bullet\overrightarrow{a} + 0 + \overrightarrow{b}\bullet\overrightarrow{b} \\ &= ||\overrightarrow{a}||^2 + ||\overrightarrow{b}||^2, \end{aligned}$$

from where the desired result follows. ❑

363 Definition The *projection* of $\overrightarrow{t}$ onto $\overrightarrow{v}$ (or the $\overrightarrow{v}$-component of $\overrightarrow{t}$) is the vector

$$\mathbf{proj}_{\overrightarrow{v}}^{\overrightarrow{t}} = (\cos\left(\widehat{\overrightarrow{t}, \overrightarrow{v}}\right))\left|\left|\overrightarrow{t}\right|\right|\frac{1}{\left|\left|\overrightarrow{v}\right|\right|}\overrightarrow{v},$$

where $(\widehat{\overrightarrow{v}, \overrightarrow{t}}) \in [0; \pi]$ is the convex angle between $\overrightarrow{v}$ and $\overrightarrow{t}$ read in the positive sense.

☞ *Given two vectors $\overrightarrow{t}$ and vector $\overrightarrow{v} \neq \overrightarrow{0}$, find bi-point representatives of them having a common tail and join them together at their tails. The projection of $\overrightarrow{t}$ onto $\overrightarrow{v}$ is the "shadow" of $\overrightarrow{t}$ in the direction of $\overrightarrow{v}$. To obtain $\mathbf{proj}_{\overrightarrow{v}}^{\overrightarrow{t}}$ we prolong $\overrightarrow{v}$ if necessary and drop a perpendicular line to it from the head of $\overrightarrow{t}$. The projection is the portion between the common tails of the vectors and the point where this perpendicular meets $\overrightarrow{t}$. See figure 8.20.*

Figure 8.20: Vector Projections.

364 Corollary Let $\overrightarrow{a} \neq \overrightarrow{0}$. Then

$$\mathbf{proj}_{\overrightarrow{a}}^{\overrightarrow{x}} = \cos\left(\widehat{\overrightarrow{x}, \overrightarrow{a}}\right)\left|\left|\overrightarrow{x}\right|\right|\frac{1}{\left|\left|\overrightarrow{a}\right|\right|}\overrightarrow{a} = \frac{\overrightarrow{x}\bullet\overrightarrow{a}}{\left|\left|\overrightarrow{a}\right|\right|^2}\overrightarrow{a}.$$

365 Theorem Let $\overrightarrow{a} \in \mathbb{R}^2 \setminus \{\overrightarrow{0}\}$. Then any $\overrightarrow{x} \in \mathbb{R}^2$ can be decomposed as

$$\overrightarrow{x} = \overrightarrow{u} + \overrightarrow{v},$$

where $\overrightarrow{u} \in \mathbb{R}\overrightarrow{a}$ and $\overrightarrow{v} \in \overrightarrow{a}^{\perp}$.

Proof: *We know that $\mathbf{proj}_{\overrightarrow{a}}^{\overrightarrow{x}}$ is parallel to $\overrightarrow{a}$, so we take $\overrightarrow{u} = \mathbf{proj}_{\overrightarrow{a}}^{\overrightarrow{x}}$. This means that we must then take $\overrightarrow{v} = \overrightarrow{x} - \mathbf{proj}_{\overrightarrow{a}}^{\overrightarrow{x}}$. We must demonstrate that $\overrightarrow{v}$ is indeed perpendicular to $\overrightarrow{a}$. But this is clear, as*

$$\begin{aligned}\overrightarrow{a}\bullet\overrightarrow{v} &= \overrightarrow{a}\bullet\overrightarrow{x} - \overrightarrow{a}\bullet\mathbf{proj}_{\overrightarrow{a}}^{\overrightarrow{x}} \\ &= \overrightarrow{a}\bullet\overrightarrow{x} - \overrightarrow{a}\bullet\frac{\overrightarrow{x}\bullet\overrightarrow{a}}{\left|\left|\overrightarrow{a}\right|\right|^2}\overrightarrow{a} \\ &= \overrightarrow{a}\bullet\overrightarrow{x} - \overrightarrow{x}\bullet\overrightarrow{a} \\ &= 0,\end{aligned}$$

completing the proof. ❑

366 Corollary Let $\overrightarrow{v} \perp \overrightarrow{w}$ be non-zero vectors in $\mathbb{R}^2$. Then any vector $\overrightarrow{a} \in \mathbb{R}^2$ has a unique representation as a linear combination of $\overrightarrow{v}, \overrightarrow{w}$,

$$\overrightarrow{a} = s\overrightarrow{v} + t\overrightarrow{w}, \quad (s, t) \in \mathbb{R}^2.$$

Proof: *By Theorem 365, there exists a decomposition*

$$\vec{a} = s\vec{v} + s'\vec{v}',$$

where $\vec{v}'$ *is orthogonal to* $\vec{v}$. *But then* $\vec{v}' || \vec{w}$ *and hence there exists* $\alpha \in \mathbb{R}$ *with* $\vec{v}' = \alpha\vec{w}$. *Taking* $t = s'\alpha$ *we achieve the decomposition*

$$\vec{a} = s\vec{v} + t\vec{w}.$$

To prove uniqueness, assume

$$s\vec{v} + t\vec{w} = \vec{a} = p\vec{v} + q\vec{w}.$$

Then $(s - p)\vec{v} = (q - t)\vec{w}$. *We must have* $s = p$ *and* $q = t$ *since otherwise* $\vec{v}$ *would be parallel to* $\vec{w}$. *This completes the proof.* ❑

367 Corollary Let $\vec{p}, \vec{q}$ be non-zero, non-parallel vectors in $\mathbb{R}^2$. Then any vector $\vec{a} \in \mathbb{R}^2$ has a unique representation as a linear combination of $\vec{p}, \vec{q}$,

$$\vec{a} = l\vec{p} + m\vec{q}, \quad (l, m) \in \mathbb{R}^2.$$

Proof: *Consider* $\vec{z} = \vec{q} - \mathbf{proj}_{\vec{p}}^{\vec{q}}$. *Clearly* $\vec{p} \perp \vec{z}$ *and so by Corollary 366, there exists unique* $(s, t) \in \mathbb{R}^2$ *such that*

$$\begin{aligned} \vec{a} &= s\vec{p} + t\vec{z} \\ &= s\vec{p} - t\mathbf{proj}_{\vec{p}}^{\vec{q}} + t\vec{q} \\ &= \left(s - t\frac{\vec{q}\bullet\vec{p}}{||\vec{p}||^2}\right)\vec{p} + t\vec{q}, \end{aligned}$$

establishing the result upon choosing $l = s - t\frac{\vec{q}\bullet\vec{p}}{||\vec{p}||^2}$ *and* $m = t$. ❑

368 Example Let $\vec{p} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{q} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. Write $\vec{p}$ as the sum of two vectors, one parallel to $\vec{q}$ and the other perpendicular to $\vec{q}$.

Solution: ▶ *We use Theorem 365. We know that* $\mathbf{proj}_{\vec{q}}^{\vec{p}}$ *is parallel to* $\vec{q}$, *and we find*

$$\mathbf{proj}_{\vec{q}}^{\vec{p}} = \frac{\vec{p}\bullet\vec{q}}{||\vec{q}||^2}\vec{q} = \frac{3}{5}\vec{q} = \begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \end{bmatrix}.$$

We also compute

$$\vec{p} - \mathbf{proj}_{\vec{q}}^{\vec{p}} = \begin{bmatrix} 1 - \frac{3}{5} \\ 1 - \frac{6}{5} \end{bmatrix} = \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \end{bmatrix}.$$

Observe that

$$\begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \end{bmatrix} \bullet \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \end{bmatrix} = \frac{6}{25} - \frac{6}{25} = 0,$$

and the desired decomposition is

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{5} \\ \frac{6}{5} \end{bmatrix} + \begin{bmatrix} \frac{2}{5} \\ -\frac{1}{5} \end{bmatrix}.$$

◀

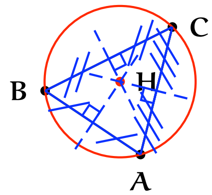


Figure 8.21: Orthocentre.

369 Example Prove that the altitudes of a triangle $\triangle ABC$ on the plane are concurrent. This point is called the *orthocentre* of the triangle.

Solution: ▶ *Put* $\vec{a} = \overrightarrow{OA}, \vec{b} = \overrightarrow{OB}, \vec{c} = \overrightarrow{OC}$. *First observe that for any* $\vec{x}$, *we have, upon expanding,*

$$(\vec{x} - \vec{a})\bullet(\vec{b} - \vec{c}) + (\vec{x} - \vec{b})\bullet(\vec{c} - \vec{a}) + (\vec{x} - \vec{c})\bullet(\vec{a} - \vec{b}) = 0. \tag{8.11}$$

Let H *be the point of intersection of the altitude from* A *and the altitude from* B. *Then*

$$0 = \overrightarrow{AH}\bullet\overrightarrow{CB} = (\overrightarrow{OH} - \overrightarrow{OA})\bullet(\overrightarrow{OB} - \overrightarrow{OC}) = (\overrightarrow{OH} - \vec{a})\bullet(\vec{b} - \vec{c}), \tag{8.12}$$

and

$$0 = \overrightarrow{BH}\bullet\overrightarrow{AC} = (\overrightarrow{OH} - \overrightarrow{OB})\bullet(\overrightarrow{OC} - \overrightarrow{OA}) = (\overrightarrow{OH} - \vec{b})\bullet(\vec{c} - \vec{a}). \tag{8.13}$$

Putting $\vec{x} = \overrightarrow{OH}$ *in (8.11) and subtracting from it (8.12) and (8.13), we gather that*

$$0 = (\overrightarrow{OH} - \vec{c})\bullet(\vec{a} - \vec{b}) = \overrightarrow{CH}\bullet\overrightarrow{AB},$$

which gives the result. ◀

Homework

Problem 8.3.1 Determine the value of a so that $\begin{bmatrix} a \\ 1-a \end{bmatrix}$ be perpendicular to $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Problem 8.3.2 Demonstrate that

$$(\vec{b} + \vec{c} = \vec{0}) \wedge (||\vec{a}|| = ||\vec{b}||) \iff (\vec{a} - \vec{b})\bullet(\vec{a} - \vec{c}) = 0.$$

Problem 8.3.3 Let $\vec{p} = \begin{bmatrix} 4 \\ 5 \end{bmatrix}$, $\vec{r} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, $\vec{s} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Write $\vec{p}$ as the sum of two vectors, one parallel to $\vec{r}$ and the other parallel to $\vec{s}$.

Problem 8.3.4 Prove that

$$||\vec{a}||^2 = (\vec{a}\bullet\vec{i})^2 + (\vec{a}\bullet\vec{j})^2.$$

Problem 8.3.5 Let $\vec{a} \neq \vec{0} \neq \vec{b}$ be vectors in $\mathbb{R}^2$ such that $\vec{a}\bullet\vec{b} = 0$. Prove that

$$\alpha\vec{a} + \beta\vec{b} = \vec{0} \implies \alpha = \beta = 0.$$

Problem 8.3.6 Let $(\vec{x}, \vec{y}) \in (\mathbb{R}^2)^2$ with $||\vec{x}|| = \frac{3}{2}||\vec{y}||$. Shew that $2\vec{x} + 3\vec{y}$ and $2\vec{x} - 3\vec{y}$ are perpendicular.

Problem 8.3.7 Let $\vec{a}, \vec{b}$ be fixed vectors in $\mathbb{R}^2$. Prove that if

$$\forall \vec{v} \in \mathbb{R}^2, \vec{v} \bullet \vec{a} = \vec{v} \bullet \vec{b},$$

then $\vec{a} = \vec{b}$.

Problem 8.3.8 Let $(\vec{a}, \vec{b}) \in (\mathbb{R}^2)^2$. Prove that

$$\left|\left|\vec{a} + \vec{b}\right|\right|^2 + \left|\left|\vec{a} - \vec{b}\right|\right|^2 = 2\left|\left|\vec{a}\right|\right|^2 + 2\left|\left|\vec{b}\right|\right|^2.$$

Problem 8.3.9 Let $\vec{u}, \vec{v}$ be vectors in $\mathbb{R}^2$. Prove the *polarisation identity:*

$$\vec{u} \bullet \vec{v} = \frac{1}{4}\left(||\vec{u} + \vec{v}||^2 - ||\vec{u} - \vec{v}||^2\right).$$

Problem 8.3.10 Let $\vec{x}, \vec{a}$ be non-zero vectors in $\mathbb{R}^2$. Prove that

$$\mathbf{proj}_{\vec{a}}^{\mathbf{proj}_{\vec{x}}^{\vec{a}}} = \alpha \vec{a},$$

with $0 \leq \alpha \leq 1$.

Problem 8.3.11 Let $(\lambda, \vec{a}) \in \mathbb{R} \times \mathbb{R}^2$ be fixed. Solve the equation

$$\vec{a} \bullet \vec{x} = \lambda$$

for $\vec{x} \in \mathbb{R}^2$.

8.4 Lines on the Plane

370 Definition Three points A, B, and C are *collinear* if they lie on the same line.

It is clear that the points A, B, and C are collinear if and only if $\overrightarrow{AB}$ is parallel to $\overrightarrow{AC}$. Thus we have the following definition.

371 Definition The parametric equation with parameter $t \in \mathbb{R}$ of the straight line passing through the point $P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}$ in the direction of the vector $\vec{v} \neq \vec{0}$ is

$$\begin{bmatrix} x - p_1 \\ y - p_2 \end{bmatrix} = t\vec{v}.$$

If $\vec{r} = \begin{bmatrix} x \\ y \end{bmatrix}$, then the equation of the line can be written in the form

$$\vec{r} - \vec{p} = t\vec{v}. \tag{8.14}$$

The *Cartesian equation of a line* is an equation of the form $ax + by = c$, where $a^2 + b^2 \neq 0$. We write (AB) for the line passing through the points A and B.

372 Theorem Let $\vec{v} \neq \vec{0}$ and let $\vec{n} \perp \vec{v}$. An alternative form for the equation of the line $\vec{r} - \vec{p} = t\vec{v}$ is

$$(\vec{r} - \vec{p}) \bullet \vec{n} = 0.$$

Moreover, the vector $\begin{bmatrix} a \\ b \end{bmatrix}$ is perpendicular to the line with Cartesian equation $ax + by = c$.

Proof: *The first part follows at once by observing that* $\vec{v} \bullet \vec{n} = 0$ *and taking dot products to both sides of* 8.14. *For the second part observe that at least one of* a *and* b *is* $\neq 0$. *First assume that*

$a \neq 0$. Then we can put $y = t$ and $x = -\frac{b}{a}t + \frac{c}{a}$ and the parametric equation of this line is

$$\begin{bmatrix} x - \frac{c}{a} \\ y \end{bmatrix} = t \begin{bmatrix} -\frac{b}{a} \\ 1 \end{bmatrix},$$

and we have

$$\begin{bmatrix} -\frac{b}{a} \\ 1 \end{bmatrix} \bullet \begin{bmatrix} a \\ b \end{bmatrix} = -\frac{b}{a} \cdot a + b = 0.$$

Similarly if $b \neq 0$ we can put $x = t$ and $y = -\frac{a}{b}t + \frac{c}{b}$ and the parametric equation of this line is

$$\begin{bmatrix} x \\ y - \frac{c}{b} \end{bmatrix} = t \begin{bmatrix} 1 \\ -\frac{a}{b} \end{bmatrix},$$

and we have

$$\begin{bmatrix} 1 \\ -\frac{a}{b} \end{bmatrix} \bullet \begin{bmatrix} a \\ b \end{bmatrix} = a - \frac{a}{b} \cdot b = 0,$$

proving the theorem in this case. ❑

☞ *The vector* $\begin{bmatrix} \frac{a}{\sqrt{a^2+b^2}} \\ \frac{b}{\sqrt{a^2+b^2}} \end{bmatrix}$ *has norm 1 and is orthogonal to the line $ax + by = c$.*

373 Example The equation of the line passing through $A = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ and in the direction of $\overrightarrow{v} = \begin{bmatrix} -4 \\ 5 \end{bmatrix}$ is

$$\begin{bmatrix} x - 2 \\ y - 3 \end{bmatrix} = \lambda \begin{bmatrix} -4 \\ 5 \end{bmatrix}.$$

374 Example Find the equation of the line passing through $A = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ and $B = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$.

Solution: ► *The direction of this line is that of*

$$\overrightarrow{AB} = \begin{bmatrix} -2 - (-1) \\ 3 - 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$

The equation is thus

$$\begin{bmatrix} x + 1 \\ y - 1 \end{bmatrix} = \lambda \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \lambda \in \mathbb{R}.$$

◀

375 Example Suppose that $(m, b) \in \mathbb{R}^2$. Write the Cartesian equation of the line $y = mx + b$ in parametric form.

Solution: ▶ *Here is a way. Put* $x = t$. *Then* $y = mt + b$ *and so the desired parametric form is*

$$\begin{bmatrix} x \\ y - b \end{bmatrix} = t \begin{bmatrix} 1 \\ m \end{bmatrix}.$$

◀

376 Example Let $(m_1, m_2, b_1, b_2) \in \mathbb{R}^4, m_1 m_2 \neq 0$. Consider the lines $L_1 : y = m_1 x + b_1$ and $L_2 : y = m_2 x + b_2$. By translating this problem in the language of vectors in $\mathbb{R}^2$, shew that $L_1 \perp L_2$ if and only if $m_1 m_2 = -1$.

Solution: ▶ *The parametric equations of the lines are*

$$L_1 : \begin{bmatrix} x \\ y - b_1 \end{bmatrix} = s \begin{bmatrix} 1 \\ m_1 \end{bmatrix}, \quad L_2 : \begin{bmatrix} x \\ y - b_2 \end{bmatrix} = t \begin{bmatrix} 1 \\ m_2 \end{bmatrix}.$$

Put $\overrightarrow{v} = \begin{bmatrix} 1 \\ m_1 \end{bmatrix}$ *and* $\overrightarrow{w} = \begin{bmatrix} 1 \\ m_2 \end{bmatrix}$. *Since the lines are perpendicular we must have* $\overrightarrow{v} \bullet \overrightarrow{w} = 0$, *which yields*

$$0 = \overrightarrow{v} \bullet \overrightarrow{w} = 1(1) + m_1(m_2) \implies m_1 m_2 = -1.$$

◀

377 Theorem (Distance Between a Point and a Line) Let $(\overrightarrow{r} - \overrightarrow{a}) \bullet \overrightarrow{n} = 0$ be a line passing through the point A and perpendicular to vector $\overrightarrow{n}$. If B is not a point on the line, then the distance from B to the line is

$$\frac{\left|(\overrightarrow{a} - \overrightarrow{b}) \bullet \overrightarrow{n}\right|}{\left|\left|\overrightarrow{n}\right|\right|}.$$

If the line has Cartesian equation $ax + by = c$, then this distance is

$$\frac{|ab_1 + bb_2 - c|}{\sqrt{a^2 + b^2}}.$$

Proof: *Let* R_0 *be the point on the line that is nearest to* B. *Then* $\overrightarrow{BR_0} = \overrightarrow{r_0} - \overrightarrow{b}$ *is orthogonal to the line, and the distance we seek is*

$$||\mathbf{proj}_{\overrightarrow{n}}^{\overrightarrow{\overrightarrow{r_0} - \overrightarrow{b}}}|| = \left|\left|\frac{(\overrightarrow{r_0} - \overrightarrow{b}) \bullet \overrightarrow{n}}{||\overrightarrow{n}||^2} \overrightarrow{n}\right|\right| = \frac{|(\overrightarrow{r_0} - \overrightarrow{b}) \bullet \overrightarrow{n}|}{||\overrightarrow{n}||}.$$

Since R_0 *is on the line,* $\overrightarrow{r_0} \bullet \overrightarrow{n} = \overrightarrow{a} \bullet \overrightarrow{n}$, *and so*

$$||\mathbf{proj}_{\overrightarrow{n}}^{\overrightarrow{\overrightarrow{r_0} - \overrightarrow{b}}}|| = \frac{|\overrightarrow{r_0} \bullet \overrightarrow{n} - \overrightarrow{b} \bullet \overrightarrow{n}|}{||\overrightarrow{n}||} = \frac{|\overrightarrow{a} \bullet \overrightarrow{n} - \overrightarrow{b} \bullet \overrightarrow{n}|}{||\overrightarrow{n}||} = \frac{|(\overrightarrow{a} - \overrightarrow{b}) \bullet \overrightarrow{n}|}{||\overrightarrow{n}||},$$

as we wanted to shew.

If the line has Cartesian equation $ax+by=c$, then at least one of a and b is $\neq 0$. Let us suppose $a\neq 0$, as the argument when $a=0$ and $b\neq 0$ is similar. Then $ax+by=c$ is equivalent to

$$\left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \frac{c}{a} \\ 0 \end{bmatrix}\right) \bullet \begin{bmatrix} a \\ b \end{bmatrix} = 0.$$

We use the result obtained above with $\overrightarrow{a} = \begin{bmatrix} \frac{c}{a} \\ 0 \end{bmatrix}$, $\overrightarrow{n} = \begin{bmatrix} a \\ b \end{bmatrix}$, *and* $B = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$. *Then* $||\overrightarrow{n}|| = \sqrt{a^2+b^2}$ *and*

$$|(\overrightarrow{a}-\overrightarrow{b})\bullet\overrightarrow{n}| = \left|\begin{bmatrix} \frac{c}{a}-b_1 \\ -b_2 \end{bmatrix} \bullet \begin{bmatrix} a \\ b \end{bmatrix}\right| = |c-ab_1-bb_2| = |ab_1+bb_2-c|,$$

giving the result. ❑

378 Example Recall that the medians of $\triangle ABC$ are lines joining the vertices of $\triangle ABC$ with the midpoints of the side opposite the vertex. Prove that the medians of a triangle are concurrent, that is, that they pass through a common point.

☞ *This point of concurrency is called, alternatively, the* isobarycentre, centroid, *or* centre of gravity *of the triangle.*

Solution: ▸ *Let M_A, M_B, and M_C denote the midpoints of the lines opposite A, B, and C, respectively. The equation of the line passing through A and in the direction of $\overrightarrow{AM_A}$ is (with* $\overrightarrow{r} = \begin{bmatrix} x \\ y \end{bmatrix}$*)*

$$\overrightarrow{r} = \overrightarrow{OA} + r\overrightarrow{AM_A}.$$

Similarly, the equation of the line passing through B and in the direction of $\overrightarrow{BM_B}$ is

$$\overrightarrow{r} = \overrightarrow{OB} + s\overrightarrow{BM_A}.$$

These two lines must intersect at a point G inside the triangle. We will shew that $\overrightarrow{GC}$ is parallel to $\overrightarrow{CM_C}$, which means that the three points G, C, M_C are collinear.

Now, $\exists(r_0, s_0) \in \mathbb{R}^2$ such that

$$\overrightarrow{OA} + r_0\overrightarrow{AM_A} = \overrightarrow{OG} = \overrightarrow{OB} + s_0\overrightarrow{BM_B},$$

that is

$$r_0\overrightarrow{AM_A} - s_0\overrightarrow{BM_B} = \overrightarrow{OB} - \overrightarrow{OA},$$

or

$$r_0(\overrightarrow{AB} + \overrightarrow{BM_A}) - s_0(\overrightarrow{BA} + \overrightarrow{AM_B}) = \overrightarrow{AB}.$$

Since M_A and M_B are the midpoints of $[B, C]$ and $[C, A]$ respectively, we have $2\overrightarrow{BM_A} = \overrightarrow{BC}$ and $2\overrightarrow{AM_B} = \overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{BC}$. The relationship becomes

$$r_0(\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BC}) - s_0(-\overrightarrow{AB} + \frac{1}{2}\overrightarrow{AB} + \frac{1}{2}\overrightarrow{BC}) = \overrightarrow{AB},$$

$$(r_0 + \frac{s_0}{2} - 1)\overrightarrow{AB} = (-\frac{r_0}{2} + \frac{s_0}{2})\overrightarrow{BC}.$$

We must have

$$r_0 + \frac{s_0}{2} - 1 = 0,$$
$$-\frac{r_0}{2} + \frac{s_0}{2} = 0,$$

since otherwise the vectors $\overrightarrow{AB}$ and $\overrightarrow{BC}$ would be parallel, and the triangle would be degenerate. Solving, we find $s_0 = r_0 = \frac{2}{3}$. Thus we have $\overrightarrow{OA} + \frac{2}{3}\overrightarrow{AM_A} = \overrightarrow{OG}$, or $\overrightarrow{AG} = \frac{2}{3}\overrightarrow{AM_A}$, and similarly, $\overrightarrow{BG} = \frac{2}{3}\overrightarrow{BM_B}$.

From $\overrightarrow{AG} = \frac{2}{3}\overrightarrow{AM_A}$, we deduce $\overrightarrow{AG} = 2\overrightarrow{GM_A}$. Since M_A is the midpoint of $[B, C]$, we have $\overrightarrow{GB} + \overrightarrow{GC} = 2\overrightarrow{GM_A} = \overrightarrow{AG}$, which is equivalent to

$$\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = \overrightarrow{0}.$$

As M_C is the midpoint of $[A, B]$ we have $\overrightarrow{GA} + \overrightarrow{GB} = 2\overrightarrow{GM_C}$. Thus

$$\overrightarrow{0} = \overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = 2\overrightarrow{GM_C} + \overrightarrow{GC}.$$

This means that $\overrightarrow{GC} = -2\overrightarrow{GM_C}$, that is, that they are parallel, and so the points G, C and M_C all lie on the same line. This achieves the desired result. ◄

☞ *The centroid of $\triangle ABC$ satisfies thus*

$$\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = \overrightarrow{0},$$

and divides the medians on the ratio $2 : 1$, reckoning from a vertex.

Homework

Problem 8.4.1 Find the angle between the lines $2x - y = 1$ and $x - 3y = 1$.

Problem 8.4.2 Find the equation of the line passing through $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and in a direction perpendicular to $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Problem 8.4.3 $\triangle ABC$ has centroid G, and $\triangle A'B'C'$ satisfies

$$\overrightarrow{AA'} + \overrightarrow{BB'} + \overrightarrow{CC'} = \overrightarrow{0}.$$

Prove that G is also the centroid of $\triangle A'B'C'$.

Problem 8.4.4 Let $ABCD$ be a trapezoid, with bases $[A, B]$ and $[C, D]$. The lines (AC) and (BD) meet at E and the lines (AD) and (BC) meet at F. Prove that the line (EF) passes through the midpoints of $[A, B]$ and $[C, D]$ by proving the following steps.

❶ Let I be the midpoint of $[A, B]$ and let J be the point of intersection of the lines (FI) and (DC). Prove that J is the midpoint of $[C, D]$. Deduce that F, I, J are collinear.

❷ Prove that E, I, J are collinear.

Problem 8.4.5 Let $ABCD$ be a parallelogram.

❶ Let E and F be points such that

$$\overrightarrow{AE} = \frac{1}{4}\overrightarrow{AC} \quad \textbf{and} \quad \overrightarrow{AF} = \frac{3}{4}\overrightarrow{AC}.$$

Demonstrate that the lines (BE) and (DF) are parallel.

❷ Let I be the midpoint of $[A, D]$ and J be the midpoint of $[B, C]$. Demonstrate that the lines (AB) and (IJ) are parallel. What type of quadrilateral is $IEJF$?

Problem 8.4.6 $ABCD$ is a parallelogram; point I is the midpoint of $[A, B]$. Point E is defined by the relation $\overrightarrow{IE} = \frac{1}{3}\overrightarrow{ID}$. Prove that

$$\overrightarrow{AE} = \frac{1}{3}\left(\overrightarrow{AB} + \overrightarrow{AD}\right)$$

and prove that the points A, C, E are collinear.

Problem 8.4.7 Put $\overrightarrow{OA} = \overrightarrow{a}$, $\overrightarrow{OB} = \overrightarrow{b}$, $\overrightarrow{OC} = \overrightarrow{c}$. Prove that A, B, C are collinear if and only if there exist real numbers α, β, γ, not all zero, such that

$$\alpha\overrightarrow{a} + \beta\overrightarrow{b} + \gamma\overrightarrow{c} = \overrightarrow{0}, \quad \alpha + \beta + \gamma = 0.$$

Problem 8.4.8 Prove Desargues' Theorem: If $\triangle ABC$ and $\triangle A'B'C'$ (not necessarily in the same plane) are so positioned that (AA'), (BB'), (CC') all pass through the same point V and if (BC) and $(B'C')$ meet at L, (CA) and $(C'A')$ meet at M, and (AB) and $(A'B')$ meet at N, then L, M, N are collinear.

8.5 Vectors in $\mathbb{R}^3$

We now extend the notions studied for $\mathbb{R}^2$ to $\mathbb{R}^3$. The rectangular coordinate form of a vector in $\mathbb{R}^3$ is

$$\overrightarrow{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

In particular, if

$$\overrightarrow{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \overrightarrow{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \overrightarrow{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

then we can write any vector $\overrightarrow{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ as a sum

$$\overrightarrow{a} = a_1\overrightarrow{i} + a_2\overrightarrow{j} + a_3\overrightarrow{k}.$$

Given $\overrightarrow{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$ and $\overrightarrow{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, their dot product is

$$\overrightarrow{a}\bullet\overrightarrow{b} = a_1b_1 + a_2b_2 + a_3b_3,$$

and

$$\left|\left|\overrightarrow{a}\right|\right| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

We also have

$$\overrightarrow{i}\bullet\overrightarrow{j} = \overrightarrow{j}\bullet\overrightarrow{k} = \overrightarrow{k}\bullet\overrightarrow{i} = 0,$$

and

$$\left|\left|\overrightarrow{i}\right|\right| = \left|\left|\overrightarrow{j}\right|\right| = \left|\left|\overrightarrow{k}\right|\right| = 1.$$

379 Definition A system of unit vectors $\overrightarrow{i}, \overrightarrow{j}, \overrightarrow{k}$ is *right-handed* if the shortest-route rotation which brings $\overrightarrow{i}$ to coincide with $\overrightarrow{j}$ is performed in a counter-clockwise manner. It is *left-handed* if the rotation is done in a clockwise manner.

To study points in space we must first agree on the orientation that we will give our coordinate system. We will use, unless otherwise noted, a right-handed orientation, as in figure 8.22.

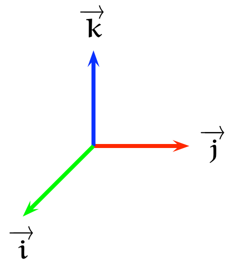


Figure 8.22: Right-handed system.

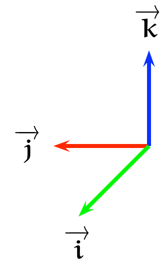


Figure 8.23: Left-handed system.

☞ *The analogues of the Cauchy-Bunyakovsky-Schwarz and the Triangle Inequality also hold in* $\mathbb{R}^3$.

We now define the (standard) cross (wedge) product in $\mathbb{R}^3$ as a product satisfying the following properties.

380 Definition Let $(\overrightarrow{x}, \overrightarrow{y}, \overrightarrow{z}, \alpha) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}$. The wedge product $\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is a closed binary operation satisfying

CP1 **Anti-commutativity:**

$$\overrightarrow{x} \times \overrightarrow{y} = -(\overrightarrow{y} \times \overrightarrow{x}) \tag{8.15}$$

CP2 **Bilinearity:**

$$(\overrightarrow{x} + \overrightarrow{z}) \times \overrightarrow{y} = \overrightarrow{x} \times \overrightarrow{y} + \overrightarrow{z} \times \overrightarrow{y}, \quad \overrightarrow{x} \times (\overrightarrow{z} + \overrightarrow{y}) = \overrightarrow{x} \times \overrightarrow{z} + \overrightarrow{x} \times \overrightarrow{y} \tag{8.16}$$

CP3 **Scalar homogeneity:**

$$(\alpha\overrightarrow{x}) \times \overrightarrow{y} = \overrightarrow{x} \times (\alpha\overrightarrow{y}) = \alpha(\overrightarrow{x} \times \overrightarrow{y}) \tag{8.17}$$

CP4

$$\overrightarrow{x} \times \overrightarrow{x} = \overrightarrow{\mathbf{0}} \tag{8.18}$$

CP5 **Right-hand Rule:**

$$\overrightarrow{i} \times \overrightarrow{j} = \overrightarrow{k}, \quad \overrightarrow{j} \times \overrightarrow{k} = \overrightarrow{i}, \quad \overrightarrow{k} \times \overrightarrow{i} = \overrightarrow{j} \tag{8.19}$$

381 Theorem Let $\overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\overrightarrow{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$ be vectors in $\mathbb{R}^3$. Then

$$\overrightarrow{x} \times \overrightarrow{y} = (x_2y_3 - x_3y_2)\overrightarrow{i} + (x_3y_1 - x_1y_3)\overrightarrow{j} + (x_1y_2 - x_2y_1)\overrightarrow{k}.$$

Proof: *Since* $\vec{i}\times\vec{i}=\vec{j}\times\vec{j}=\vec{k}\times\vec{k}=\vec{0}$ *we have*

$$\begin{aligned}(x_1\vec{i}+x_2\vec{j}+x_3\vec{k})\times(y_1\vec{i}+y_2\vec{j}+y_3\vec{k}) &= x_1y_2\vec{i}\times\vec{j}+x_1y_3\vec{i}\times\vec{k}\\ &\quad +x_2y_1\vec{j}\times\vec{i}+x_2y_3\vec{j}\times\vec{k}\\ &\quad +x_3y_1\vec{k}\times\vec{i}+x_3y_2\vec{k}\times\vec{j}\\ &= x_1y_2\vec{k}-x_1y_3\vec{j}-x_2y_1\vec{k}\\ &\quad +x_2y_3\vec{i}+x_3y_1\vec{j}-x_3y_2\vec{i},\end{aligned}$$

from where the theorem follows. ❑

382 Example Find

$$\begin{bmatrix}1\\0\\-3\end{bmatrix}\times\begin{bmatrix}0\\1\\2\end{bmatrix}.$$

Solution: ► *We have*

$$\begin{aligned}(\vec{i}-3\vec{k})\times(\vec{j}+2\vec{k}) &= \vec{i}\times\vec{j}+2\vec{i}\times\vec{k}-3\vec{k}\times\vec{j}-6\vec{k}\times\vec{k}\\ &= \vec{k}-2\vec{j}-3\vec{i}+6\vec{0}\\ &= -3\vec{i}-2\vec{j}+\vec{k}.\end{aligned}$$

Hence

$$\begin{bmatrix}1\\0\\-3\end{bmatrix}\times\begin{bmatrix}0\\1\\2\end{bmatrix}=\begin{bmatrix}-3\\-2\\1\end{bmatrix}.$$

◄

383 Theorem The cross product vector $\vec{x}\times\vec{y}$ is simultaneously perpendicular to $\vec{x}$ and $\vec{y}$.

Proof: *We will only check the first assertion, the second verification is analogous.*

$$\begin{aligned}\vec{x}\bullet(\vec{x}\times\vec{\mathbf{y}}) &= (x_1\vec{i}+x_2\vec{j}+x_3\vec{k})\bullet((x_2y_3-x_3y_2)\vec{i}\\ &\quad +(x_3y_1-x_1y_3)\vec{j}+(x_1y_2-x_2y_1)\vec{k})\\ &= x_1x_2y_3-x_1x_3y_2+x_2x_3y_1-x_2x_1y_3+x_3x_1y_2-x_3x_2y_1\\ &= 0,\end{aligned}$$

completing the proof. ❑

384 Theorem $\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{\mathbf{a}}\bullet\vec{c})\vec{b} - (\vec{\mathbf{a}}\bullet\vec{\mathbf{b}})\vec{c}$.

Proof:

$$\begin{aligned}\vec{a}\times(\vec{b}\times\vec{c}) &= (a_1\vec{i} + a_2\vec{j} + a_3\vec{k})\times((b_2c_3 - b_3c_2)\vec{i} + \\ &\quad +(b_3c_1 - b_1c_3)\vec{j} + (b_1c_2 - b_2c_1)\vec{k}) \\ &= a_1(b_3c_1 - b_1c_3)\vec{k} - a_1(b_1c_2 - b_2c_1)\vec{j} \\ &\quad -a_2(b_2c_3 - b_3c_2)\vec{k} + a_2(b_1c_2 - b_2c_1)\vec{i} \\ &\quad +a_3(b_2c_3 - b_3c_2)\vec{j} - a_3(b_3c_1 - b_1c_3)\vec{i} \\ &= (a_1c_1 + a_2c_2 + a_3c_3)(b_1\vec{i} + b_2\vec{j} + b_3\vec{i}) \\ &\quad +(-a_1b_1 - a_2b_2 - a_3b_3)(c_1\vec{i} + c_2\vec{j} + c_3\vec{i}) \\ &= (\vec{\mathbf{a}}\bullet\vec{c})\vec{b} - (\vec{\mathbf{a}}\bullet\vec{\mathbf{b}})\vec{c},\end{aligned}$$

completing the proof. ❑

385 Theorem (Jacobi's Identity)

$$\vec{a}\times(\vec{b}\times\vec{c}) + \vec{b}\times(\vec{c}\times\vec{a}) + \vec{c}\times(\vec{a}\times\vec{b}) = \vec{0}.$$

Proof: *From Theorem 384 we have*

$$\begin{aligned}\vec{a}\times(\vec{b}\times\vec{c}) &= (\vec{\mathbf{a}}\bullet\vec{c})\vec{b} - (\vec{\mathbf{a}}\bullet\vec{\mathbf{b}})\vec{c}, \\ \vec{b}\times(\vec{c}\times\vec{a}) &= (\vec{b}\bullet\vec{a})\vec{c} - (\vec{b}\bullet\vec{c})\vec{a}, \\ \vec{c}\times(\vec{a}\times\vec{b}) &= (\vec{c}\bullet\vec{b})\vec{a} - (\vec{c}\bullet\vec{a})\vec{b},\end{aligned}$$

and adding yields the result. ❑

386 Theorem Let $\widehat{(\vec{x},\vec{y})} \in [0;\pi[$ be the convex angle between two vectors $\vec{x}$ and $\vec{y}$. Then

$$||\vec{x}\times\vec{y}|| = ||\vec{x}||||\vec{y}||\,\mathbf{sin}\,\widehat{(\vec{x},\vec{y})}.$$

Proof: *We have*

$$\begin{aligned}||\vec{x}\times\vec{y}||^2 &= (x_2y_3 - x_3y_2)^2 + (x_3y_1 - x_1y_3)^2 + (x_1y_2 - x_2y_1)^2 \\ &= x_2^2y_3^2 - 2x_2y_3x_3y_2 + x_3^2y_2^2 + x_3^2y_1^2 - 2x_3y_1x_1y_3 + \\ &\quad +x_1^2y_3^2 + x_1^2y_2^2 - 2x_1y_2x_2y_1 + x_2^2y_1^2 \\ &= (x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) - (x_1y_1 + x_2y_2 + x_3y_3)^2 \\ &= ||\vec{x}||^2||\vec{\mathbf{y}}||^2 - (\vec{\mathbf{x}}\bullet\vec{\mathbf{y}})^2 \\ &= ||\vec{x}||^2||\vec{\mathbf{y}}||^2 - ||\vec{x}||^2||\vec{\mathbf{y}}||^2\,\mathbf{cos}^2\,\widehat{(\vec{x},\vec{y})} \\ &= ||\vec{x}||^2||\vec{\mathbf{y}}||^2\,\mathbf{sin}^2\,\widehat{(\vec{x},\vec{y})},\end{aligned}$$

whence the theorem follows. The Theorem is illustrated in Figure 8.24. Geometrically it means that the area of the parallelogram generated by joining $\vec{x}$ and $\vec{y}$ at their heads is $\left|\left|\vec{x}\times\vec{y}\right|\right|$. ❑

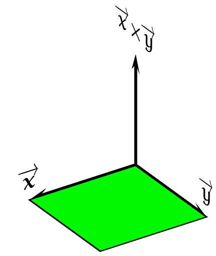


Figure 8.24: Theorem 386.

The following corollaries are now obvious.

387 Corollary Two non-zero vectors $\vec{x}, \vec{y}$ satisfy $\vec{x}\times\vec{y} = \vec{\mathbf{0}}$ if and only if they are parallel.

388 Corollary (Lagrange's Identity)

$$||\vec{x}\times\vec{\mathbf{y}}||^2 = ||\mathbf{x}||^2||\mathbf{y}||^2 - (\vec{\mathbf{x}}\bullet\vec{\mathbf{y}})^2.$$

389 Example Let $\vec{x} \in \mathbb{R}^3$, $||\mathbf{x}|| = 1$. Find

$$||\vec{x}\times\vec{i}||^2 + ||\vec{x}\times\vec{j}||^2 + ||\vec{x}\times\vec{k}||^2.$$

Solution: ► *By Lagrange's Identity,*

$$||\vec{x}\times\vec{i}||^2 = \left|\left|\vec{x}\right|\right|^2\left|\left|\vec{i}\right|\right|^2 - (\vec{x}\bullet\vec{i})^2 = 1 - (\vec{\mathbf{x}}\bullet\vec{i})^2,$$

$$||\vec{x}\times\vec{k}||^2 = \left|\left|\vec{x}\right|\right|^2\left|\left|\vec{j}\right|\right|^2 - (\vec{x}\bullet\vec{j})^2 = 1 - (\vec{\mathbf{x}}\bullet\vec{j})^2,$$

$$||\vec{x}\times\vec{j}||^2 = \left|\left|\vec{x}\right|\right|^2\left|\left|\vec{k}\right|\right|^2 - (\vec{x}\bullet\vec{k})^2 = 1 - (\vec{x}\bullet\vec{k})^2,$$

and since $(\vec{x}\bullet\vec{i})^2 + (\vec{x}\bullet\vec{j})^2 + (\vec{x}\bullet\vec{k})^2 = \left|\left|\vec{x}\right|\right|^2 = 1$, *the desired sum equals* $3 - 1 = 2$. ◄

Problem 8.5.1 Consider a tetrahedron $ABCS$. [A] Find $\overrightarrow{AB} + \overrightarrow{BC} + \overrightarrow{CS}$. [B] Find $\overrightarrow{AC} + \overrightarrow{CS} + \overrightarrow{SA} + \overrightarrow{AB}$.

Problem 8.5.2 Find a vector simultaneously perpendicular to $\begin{bmatrix}1\\1\\1\end{bmatrix}$ and $\begin{bmatrix}1\\1\\0\end{bmatrix}$ and having norm 3.

Problem 8.5.3 Find the area of the triangle whose vertices are at $\mathbf{P} = \begin{pmatrix}0\\0\\1\end{pmatrix}$, $\mathbf{Q} = \begin{pmatrix}0\\1\\0\end{pmatrix}$, and $\mathbf{R} = \begin{pmatrix}1\\0\\0\end{pmatrix}$.

Problem 8.5.4 Prove or disprove! The cross product is associative.

Problem 8.5.5 Prove that $\vec{x}\times\vec{x} = \vec{0}$ follows from the anti-commutativity of the cross product.

Problem 8.5.6 Expand the product $(\vec{a} - \vec{b})\times(\vec{a} + \vec{b})$.

Problem 8.5.7 The vectors $\vec{a}, \vec{b}$ are constant vectors. Solve the equation $\vec{a}\times(\vec{x}\times\vec{b}) = \vec{b}\times(\vec{x}\times\vec{a})$.

Problem 8.5.8 The vectors $\vec{a}, \vec{b}, \vec{c}$ are constant vectors. Solve the system of equations

$$2\vec{x} + \vec{y}\times\vec{a} = \vec{b}, \quad 3\vec{y} + \vec{x}\times\vec{a} = \vec{c},$$

Problem 8.5.9 Prove that there do not exist three unit vectors in $\mathbb{R}^3$ such that the angle between any two of them be $> \frac{2\pi}{3}$.

Problem 8.5.10 Let $\overrightarrow{a} \in \mathbb{R}^3$ be a fixed vector. Demonstrate that

$$X = \{\overrightarrow{x} \in \mathbb{R}^3 : \overrightarrow{a} \times \overrightarrow{x} = \overrightarrow{0}\}$$

is a subspace of $\mathbb{R}^3$.

Problem 8.5.11 Let $(\overrightarrow{a}, \overrightarrow{b}) \in (\mathbb{R}^3)^2$ and assume that $\overrightarrow{\mathbf{a}} \bullet \overrightarrow{\mathbf{b}} = 0$ and that $\overrightarrow{a}$ and $\overrightarrow{b}$ are linearly independent. Prove that $\overrightarrow{a}, \overrightarrow{b}, \overrightarrow{a} \times \overrightarrow{b}$ are linearly independent.

Problem 8.5.12 Let $(\overrightarrow{a}, \overrightarrow{b}) \in \mathbb{R}^3 \times \mathbb{R}^3$ be fixed. Solve the equation

$$\overrightarrow{a} \times \overrightarrow{\mathbf{x}} = \overrightarrow{b},$$

for $\overrightarrow{x}$.

Problem 8.5.13 Let $\overrightarrow{h}, \overrightarrow{k}$ be fixed vectors in $\mathbb{R}^3$. Prove that

$$L : \begin{array}{ccc} \mathbb{R}^3 \times \mathbb{R}^3 & \to & \mathbb{R}^3 \\ (\overrightarrow{x}, \overrightarrow{y}) & \mapsto & \overrightarrow{x} \times \overrightarrow{k} + \overrightarrow{h} \times \overrightarrow{y} \end{array}$$

is a linear transformation.

8.6 Planes and Lines in $\mathbb{R}^3$

390 Definition If bi-point representatives of a family of vectors in $\mathbb{R}^3$ lie on the same plane, we will say that the vectors are *coplanar* or parallel to the plane.

391 Lemma Let $\overrightarrow{v}, \overrightarrow{w}$ in $\mathbb{R}^3$ be non-parallel vectors. Then every vector $\overrightarrow{u}$ of the form

$$\overrightarrow{u} = a\overrightarrow{v} + b\overrightarrow{w},$$

($(a, b) \in \mathbb{R}^2$ arbitrary) is coplanar with both $\overrightarrow{v}$ and $\overrightarrow{w}$. Conversely, any vector $\overrightarrow{t}$ coplanar with both $\overrightarrow{v}$ and $\overrightarrow{w}$ can be uniquely expressed in the form

$$\overrightarrow{t} = p\overrightarrow{v} + q\overrightarrow{w}.$$

Proof: *This follows at once from Corollary 367, since the operations occur on a plane, which can be identified with* $\mathbb{R}^2$. ❑

A plane is determined by three non-collinear points. Suppose that $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are non-collinear points on the same plane and that $\mathbf{R} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ is another arbitrary point on this plane. Since $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are non-collinear, $\overrightarrow{AB}$ and $\overrightarrow{AC}$, which are coplanar, are non-parallel. Since $\overrightarrow{AR}$ also lies on the plane, we have by Lemma 391, that there exist real numbers p, q with

$$\overrightarrow{AR} = p\overrightarrow{AB} + q\overrightarrow{AC}.$$

By Chasles' Rule,

$$\overrightarrow{OR} = \overrightarrow{OA} + p(\overrightarrow{OB} - \overrightarrow{OA}) + q(\overrightarrow{OC} - \overrightarrow{OA}),$$

is the equation of a plane containing the three non-collinear points $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$. By letting $\overrightarrow{r} = \overrightarrow{OR}$, $\overrightarrow{a} = \overrightarrow{OA}$, etc., we deduce that

$$\overrightarrow{r} - \overrightarrow{a} = p(\overrightarrow{b} - \overrightarrow{a}) + q(\overrightarrow{c} - \overrightarrow{a}).$$

Thus we have the following definition.

392 Definition The *parametric equation* of a plane containing the point A, and parallel to the vectors $\vec{u}$ and $\vec{v}$ is given by

$$\vec{r} - \vec{a} = p\vec{u} + q\vec{v}.$$

Componentwise this takes the form

$$x - a_1 = pu_1 + qv_1,$$

$$y - a_2 = pu_2 + qv_2,$$

$$z - a_3 = pu_3 + qv_3.$$

The *Cartesian* equation of a plane is an equation of the form $ax + by + cz = d$ with $(a, b, c, d) \in \mathbb{R}^4$ and $a^2 + b^2 + c^2 \neq 0$.

393 Example Find both the parametric equation and the Cartesian equation of the plane parallel to the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and passing through the point $\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$.

Solution: ► *The desired parametric equation is*

$$\begin{bmatrix} x \\ y+1 \\ z-2 \end{bmatrix} = s \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

This gives $s = z-2$, $t = y+1-s = y+1-z+2 = y-z+3$ *and* $x = s+t = z-2+y-z+3 = y+1$. *Hence the Cartesian equation is* $x - y = 1$. ◄

394 Theorem Let $\vec{u}$ and $\vec{v}$ be non-parallel vectors and let $\vec{r} - \vec{a} = p\vec{u} + q\vec{v}$ be the equation of the plane containing A an parallel to the vectors $\vec{u}$ and $\vec{v}$. If $\vec{n}$ is simultaneously perpendicular to $\vec{u}$ and $\vec{v}$ then

$$(\vec{r} - \vec{a})\bullet\vec{n} = 0.$$

Moreover, the vector $\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is normal to the plane with Cartesian equation $ax + by + cz = d$.

Proof: *The first part is clear, as* $\vec{u}\bullet\vec{n} = 0 = \vec{v}\bullet\vec{n}$. *For the second part, recall that at least one of* a, b, c *is non-zero. Let us assume* $a \neq 0$. *The argument is similar if one of the other letters is non-zero and* $a = 0$. *In this case we can see that*

$$x = \frac{d}{a} - \frac{b}{a}y - \frac{c}{a}z.$$

Put $y = s$ *and* $z = t$. *Then*

$$\begin{bmatrix} x - \frac{d}{a} \\ y \\ z \end{bmatrix} = s \begin{bmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{bmatrix}$$

is a parametric equation for the plane. ❑

395 Example Find once again, by appealing to Theorem 394, the Cartesian equation of the plane parallel to the vectors $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and passing through the point $\begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$.

Solution: ► *The vector* $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ *is normal to the plane. The plane has thus equation*

$$\begin{bmatrix} x \\ y+1 \\ z-2 \end{bmatrix} \bullet \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = 0 \implies -x + y + 1 = 0 \implies x - y = 1,$$

as obtained before. ◄

396 Theorem (Distance Between a Point and a Plane) Let $(\vec{r} - \vec{a})\bullet\vec{n} = 0$ be a plane passing through the point A and perpendicular to vector $\vec{n}$. If B is not a point on the plane, then the distance from B to the plane is

$$\frac{\left|(\vec{a} - \vec{b})\bullet\vec{n}\right|}{\|\vec{n}\|}.$$

Proof: *Let* R_0 *be the point on the plane that is nearest to* B. *Then* $\overrightarrow{BR_0} = \vec{r_0} - \vec{b}$ *is orthogonal to the plane, and the distance we seek is*

$$\|\mathbf{proj}_{\vec{n}}^{\overrightarrow{\vec{r_0} - \vec{b}}}\| = \left\| \frac{(\vec{r_0} - \vec{b})\bullet\vec{n}}{\|\vec{n}\|^2}\vec{n} \right\| = \frac{|(\vec{r_0} - \vec{b})\bullet\vec{n}|}{\|\vec{n}\|}.$$

Since R_0 *is on the plane,* $\vec{r_0}\bullet\vec{n} = \vec{a}\bullet\vec{n}$, *and so*

$$\|\mathbf{proj}_{\vec{n}}^{\overrightarrow{\vec{r_0} - \vec{b}}}\| = \frac{|\vec{r_0}\bullet\vec{n} - \vec{b}\bullet\vec{n}|}{\|\vec{n}\||} = \frac{|\vec{a}\bullet\vec{n} - \vec{b}\bullet\vec{n}|}{\|\vec{n}\|} = \frac{|(\vec{a} - \vec{b})\bullet\vec{n}|}{\|\vec{n}\|},$$

as we wanted to shew. ❑

☞ *Given three planes in space, they may (i) be parallel (which allows for some of them to coincide), (ii) two may be parallel and the third intersect each of the other two at a line, (iii) intersect at a line, (iv) intersect at a point.*

397 Definition The equation of a line passing through $A \in \mathbb{R}^3$ in the direction of $\vec{v} \neq \vec{0}$ is given by

$$\vec{r} - \vec{a} = t\vec{v}, \quad t \in \mathbb{R}.$$

398 Theorem Put $\overrightarrow{OA} = \vec{a}$, $\overrightarrow{OB} = \vec{b}$, and $\overrightarrow{OC} = \vec{c}$. Points $(A, B, C) \in (\mathbb{R}^3)^3$ are collinear if and only if

$$\vec{a}\times\vec{b} + \vec{b}\times\vec{c} + \vec{c}\times\vec{a} = \vec{0}.$$

Proof: *If the points A, B, C are collinear, then $\overrightarrow{AB}$ is parallel to $\overrightarrow{AC}$ and by Corollary 387, we must have*

$$(\vec{c} - \vec{a})\times(\vec{b} - \vec{a}) = \vec{0}.$$

Rearranging, gives

$$\vec{c}\times\vec{b} - \vec{c}\times\vec{a} - \vec{a}\times\vec{b} = \vec{0}.$$

Further rearranging completes the proof. ❑

399 Theorem (Distance Between a Point and a Line) Let $L : \vec{r} = \vec{a} + \lambda\vec{v}$, $\vec{v} \neq \vec{0}$, be a line and let B be a point not on L. Then the distance from B to L is given by

$$\frac{||(\vec{a} - \vec{b})\times\vec{v}||}{||\vec{v}||}.$$

Proof: *If R_0—with position vector $\vec{r_0}$—is the point on L that is at shortest distance from B then $\overrightarrow{BR_0}$ is perpendicular to the line, and so*

$$||\overrightarrow{BR_0}\times\vec{v}|| = ||\overrightarrow{BR_0}||||\vec{v}|| \sin\frac{\pi}{2} = ||\overrightarrow{BR_0}||||\vec{v}||.$$

The distance we must compute is $\left|\left|\overrightarrow{BR_0}\right|\right| = ||\vec{r_0} - \vec{b}||$, *which is then given by*

$$||\vec{r_0} - \vec{b}|| = \frac{||\overrightarrow{BR_0}\times\vec{v}||}{||\vec{v}||} = \frac{||(\vec{r_0} - \vec{b})\times\vec{v}||}{||\vec{v}||}.$$

Now, since R_0 is on the line $\exists t_0 \in \mathbb{R}$ such that $\vec{r_0} = \vec{a} + t_0\vec{v}$. Hence

$$(\vec{r_0} - \vec{b})\times\vec{v} = (\vec{a} - \vec{b})\times\vec{v},$$

giving

$$||\vec{r_0} - \vec{b}|| = \frac{||(\vec{a} - \vec{b})\times\vec{v}||}{||\vec{v}||},$$

proving the theorem. ❑

☞ *Given two lines in space, one of the following three situations might arise: (i) the lines intersect at a point, (ii) the lines are parallel, (iii) the lines are skew (one over the other, without intersecting).*

Homework

Problem 8.6.1 Find the equation of the plane passing through the points $(a, 0, a)$, $(-a, 1, 0)$, and $(0, 1, 2a)$ in $\mathbb{R}^3$.

Problem 8.6.2 Find the equation of plane containing the point $(1, 1, 1)$ and perpendicular to the line $x = 1 + t, y = -2t, z = 1 - t$.

Problem 8.6.3 Find the equation of plane containing the point $(1, -1, -1)$ and containing the line $x = 2y = 3z$.

Problem 8.6.4 Find the equation of the plane perpendicular to the line $ax = by = cz$, $abc \neq 0$ and passing through the point $(1, 1, 1)$ in $\mathbb{R}^3$.

Problem 8.6.5 Find the equation of the line perpendicular to the plane $ax + a^2y + a^3z = 0$, $a \neq 0$ and passing through the point $(0, 0, 1)$.

Problem 8.6.6 The two planes

$$x - y - z = 1, \qquad x - z = -1,$$

intersect at a line. Write the equation of this line in the form

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \overrightarrow{a} + t\overrightarrow{v}, \quad t \in \mathbb{R}.$$

Problem 8.6.7 Find the equation of the plane passing through the points $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ and parallel to the line

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}..$$

Problem 8.6.8 Points $\mathbf{a}, \mathbf{b}, \mathbf{c}$ in $\mathbb{R}^3$ are collinear and it is known that $\overrightarrow{a} \times \overrightarrow{c} = \overrightarrow{i} - 2\overrightarrow{j}$ and $\overrightarrow{a} \times \overrightarrow{b} = 2\overrightarrow{k} - 3\overrightarrow{i}$. Find $\overrightarrow{b} \times \overrightarrow{c}$.

Problem 8.6.9 Find the equation of the plane which is equidistant of the points $\begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$.

Problem 8.6.10 (Putnam Exam, 1980) Let S be the solid in three-dimensional space consisting of all points (x, y, z) satisfying the following system of six conditions:

$$x \geq 0, \quad y \geq 0, \quad z \geq 0,$$

$$x + y + z \leq 11,$$
$$2x + 4y + 3z \leq 36,$$
$$2x + 3z \leq 24.$$

Determine the number of vertices and the number of edges of S.

8.7 $\mathbb{R}^n$

As a generalisation of $\mathbb{R}^2$ and $\mathbb{R}^3$ we define $\mathbb{R}^n$ as the set of n-tuples

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} : x_i \in \mathbb{R} \right\}.$$

The dot product of two vectors in $\mathbb{R}^n$ is defined as

$$\overrightarrow{x} \bullet \overrightarrow{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \bullet \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n.$$

The norm of a vector in $\mathbb{R}^n$ is given by

$$||\overrightarrow{x}|| = \sqrt{\overrightarrow{x} \bullet \overrightarrow{x}}.$$

As in the case of $\mathbb{R}^2$ and $\mathbb{R}^3$ we have

400 Theorem (Cauchy-Bunyakovsky-Schwarz Inequality) Given $(\overrightarrow{x}, \overrightarrow{y}) \in (\mathbb{R}^n)^2$ the following inequality holds

$$|\overrightarrow{x}\bullet\overrightarrow{y}| \leq ||\overrightarrow{x}||||\overrightarrow{y}||.$$

Proof: *Put* $a = \sum_{k=1}^{n} x_k^2$, $b = \sum_{k=1}^{n} x_k y_k$, *and* $c = \sum_{k=1}^{n} y_k^2$. *Consider*

$$f(t) = \sum_{k=1}^{n} (tx_k - y_k)^2 = t^2 \sum_{k=1}^{n} x_k^2 - 2t \sum_{k=1}^{n} x_k y_k + \sum_{k=1}^{n} y_k^2 = at^2 + bt + c.$$

This is a quadratic polynomial which is non-negative for all real t, *so it must have complex roots. Its discriminant* $b^2 - 4ac$ *must be non-positive, from where we gather*

$$4\left(\sum_{k=1}^{n} x_k y_k\right)^2 \leq 4\left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right).$$

This gives

$$|\overrightarrow{x}\bullet\overrightarrow{y}|^2 \leq ||\overrightarrow{x}||^2||\overrightarrow{y}||^2$$

from where we deduce the result. ❑

401 Example Assume that $a_k, b_k, c_k, k = 1, \ldots, n$, are positive real numbers. Shew that

$$\left(\sum_{k=1}^{n} a_k b_k c_k\right)^4 \leq \left(\sum_{k=1}^{n} a_k^4\right)\left(\sum_{k=1}^{n} b_k^4\right)\left(\sum_{k=1}^{n} c_k^2\right)^2.$$

Solution: ► *Using CBS on* $\sum_{k=1}^{n} (a_k b_k) c_k$ *once we obtain*

$$\sum_{k=1}^{n} a_k b_k c_k \leq \left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2}\left(\sum_{k=1}^{n} c_k^2\right)^{1/2}.$$

Using CBS again on $\left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2}$ *we obtain*

$$\begin{aligned}\sum_{k=1}^{n} a_k b_k c_k &\leq \left(\sum_{k=1}^{n} a_k^2 b_k^2\right)^{1/2}\left(\sum_{k=1}^{n} c_k^2\right)^{1/2} \\ &\leq \left(\sum_{k=1}^{n} a_k^4\right)^{1/4}\left(\sum_{k=1}^{n} b_k^4\right)^{1/4}\left(\sum_{k=1}^{n} c_k^2\right)^{1/2},\end{aligned}$$

which gives the required inequality. ◄

402 Theorem (Triangle Inequality) Given $(\overrightarrow{x}, \overrightarrow{y}) \in (\mathbb{R}^n)^2$ the following inequality holds

$$||\overrightarrow{x} + \overrightarrow{y}|| \leq ||\overrightarrow{x}|| + ||\overrightarrow{y}||.$$

Proof: *We have*

$$\begin{aligned}||\overrightarrow{a} + \overrightarrow{b}||^2 &= (\overrightarrow{a} + \overrightarrow{b})\bullet(\overrightarrow{a} + \overrightarrow{b}) \\ &= \overrightarrow{a}\bullet\overrightarrow{a} + 2\overrightarrow{a}\bullet\overrightarrow{b} + \overrightarrow{b}\bullet\overrightarrow{b} \\ &\leq ||\overrightarrow{a}||^2 + 2||\overrightarrow{a}||||\overrightarrow{b}|| + ||\overrightarrow{b}||^2 \\ &= (||\overrightarrow{a}|| + ||\overrightarrow{b}||)^2,\end{aligned}$$

from where the desired result follows.

❑

We now consider a generalisation of the Euclidean norm. Given $p > 1$ and $\overrightarrow{x} \in \mathbb{R}^n$ we put

$$||\overrightarrow{x}||_p = \left(\sum_{k=1}^{n} |x_k|^p\right)^{1/p} \tag{8.20}$$

Clearly

$$||\overrightarrow{x}||_p \geq 0 \tag{8.21}$$

$$||\overrightarrow{x}||_p = 0 \Leftrightarrow \overrightarrow{x} = \overrightarrow{0} \tag{8.22}$$

$$||\alpha\overrightarrow{x}||_p = |\alpha|||\overrightarrow{x}||_p, \quad \alpha \in \mathbb{R} \tag{8.23}$$

We now prove analogues of the Cauchy-Bunyakovsky-Schwarz and the Triangle Inequality for $||\cdot||_p$. For this we need the following lemma.

403 Lemma (Young's Inequality) Let $p > 1$ and put $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Then for $(a, b) \in ([0; +\infty[)^2$ we have

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$

Proof: *Let $0 < k < 1$, and consider the function*

$$f : \begin{array}{ccc} [0; +\infty[& \to & \mathbb{R} \\ x & \mapsto & x^k - k(x-1) \end{array}.$$

Then $0 = f'(x) = kx^{k-1} - k \Leftrightarrow x = 1$. Since $f''(x) = k(k-1)x^{k-2} < 0$ for $0 < k < 1, x \geq 0$, $x = 1$ is a maximum point. Hence $f(x) \leq f(1)$ for $x \geq 0$, that is $x^k \leq 1 + k(x-1)$. Letting $k = \dfrac{1}{p}$ and $x = \dfrac{a^p}{b^q}$ we deduce

$$\frac{a}{b^{q/p}} \leq 1 + \frac{1}{p}\left(\frac{a^p}{b^q} - 1\right).$$

Rearranging gives

$$ab \leq b^{1+p/q} + \frac{a^p b^{1+p/q-p}}{p} - \frac{b^{1+p/q}}{p}$$

from where we obtain the inequality. ❑

The promised generalisation of the Cauchy-Bunyakovsky-Schwarz Inequality is given in the following theorem.

404 Theorem (Hölder Inequality) Given $(\overrightarrow{x}, \overrightarrow{y}) \in (\mathbb{R}^n)^2$ the following inequality holds

$$|\overrightarrow{x}\bullet\overrightarrow{y}| \leq ||\overrightarrow{x}||_p||\overrightarrow{y}||_q.$$

Proof: *If $||\overrightarrow{x}||_p = 0$ or $||\overrightarrow{y}||_q = 0$ there is nothing to prove, so assume otherwise. From the Young Inequality we have*

$$\frac{|x_k|}{||\overrightarrow{x}||_p}\frac{|y_k|}{||\overrightarrow{y}||_q} \leq \frac{|x_k|^p}{||\overrightarrow{x}||_p{}^p p} + \frac{|y_k|^q}{||\overrightarrow{y}||_q{}^q q}.$$

Adding, we deduce

$$\begin{aligned}\sum_{k=1}^n \frac{|x_k|}{||\overrightarrow{x}||_p}\frac{|y_k|}{||\overrightarrow{y}||_q} &\leq \frac{1}{||\overrightarrow{x}||_p{}^p p}\sum_{k=1}^n |x_k|^p + \frac{1}{||\overrightarrow{y}||_q{}^q q}\sum_{k=1}^n |y_k|^q \\ &= \frac{||\overrightarrow{x}||_p{}^p}{||\overrightarrow{x}||_p{}^p p} + \frac{||\overrightarrow{y}||_q{}^q}{||\overrightarrow{y}||_q{}^q q} \\ &= \frac{1}{p}+\frac{1}{q} \\ &= 1.\end{aligned}$$

This gives

$$\sum_{k=1}^n |x_k y_k| \leq ||\overrightarrow{x}||_p ||\overrightarrow{y}||_q.$$

The result follows by observing that

$$\left|\sum_{k=1}^n x_k y_k\right| \leq \sum_{k=1}^n |x_k y_k| \leq ||\overrightarrow{x}||_p ||\overrightarrow{y}||_q.$$

❑

As a generalisation of the Triangle Inequality we have

405 Theorem (Minkowski Inequality) Let $p \in]1;+\infty[$. Given $(\overrightarrow{x},\overrightarrow{y}) \in (\mathbb{R}^n)^2$ the following inequality holds

$$||\overrightarrow{x}+\overrightarrow{y}||_p \leq ||\overrightarrow{x}||_p + ||\overrightarrow{y}||_p.$$

Proof: *From the triangle inequality for real numbers 1.6*

$$|x_k+y_k|^p = |x_k+y_k||x_k+y_k|^{p-1} \leq (|x_k|+|y_k|)\,|x_k+y_k|^{p-1}.$$

Adding

$$\sum_{k=1}^n |x_k+y_k|^p \leq \sum_{k=1}^n |x_k||x_k+y_k|^{p-1} + \sum_{k=1}^n |y_k||x_k+y_k|^{p-1}. \tag{8.24}$$

By the Hölder Inequality

$$\begin{aligned}\sum_{k=1}^n |x_k||x_k+y_k|^{p-1} &\leq \left(\sum_{k=1}^n |x_k|^p\right)^{1/p}\left(\sum_{k=1}^n |x_k+y_k|^{(p-1)q}\right)^{1/q} \\ &= \left(\sum_{k=1}^n |x_k|^p\right)^{1/p}\left(\sum_{k=1}^n |x_k+y_k|^{p}\right)^{1/q} \qquad (8.25)\\ &= ||\overrightarrow{x}||_p ||\overrightarrow{x}+\overrightarrow{y}||_p^{p/q}\end{aligned}$$

In the same manner we deduce

$$\sum_{k=1}^n |y_k||x_k+y_k|^{p-1} \leq ||\overrightarrow{y}||_p ||\overrightarrow{x}+\overrightarrow{y}||_p^{p/q}. \tag{8.26}$$

Hence (8.24) gives

$$||\overrightarrow{x}+\overrightarrow{y}||_p^p = \sum_{k=1}^n |x_k+y_k|^p \leq ||\overrightarrow{x}||_p ||\overrightarrow{x}+\overrightarrow{y}||_p^{p/q} + ||\overrightarrow{y}||_p ||\overrightarrow{x}+\overrightarrow{y}||_p^{p/q},$$

from where we deduce the result. ❑

Homework

Problem 8.7.1 Prove Lagrange's identity:

$$\left(\sum_{1\leq j\leq n} a_j b_j\right)^2 = \left(\sum_{1\leq j\leq n} a_j^2\right)\left(\sum_{1\leq j\leq n} b_j^2\right) - \sum_{1\leq k<j\leq n}(a_k b_j - a_j b_k)^2$$

and then deduce the CBS Inequality in $\mathbb{R}^n$.

Problem 8.7.2 Let $\overrightarrow{a}_i \in \mathbb{R}^n$ for $1 \leq i \leq n$ be unit vectors with $\sum_{i=1}^n \overrightarrow{a_i} = \overrightarrow{0}$. Prove that $\sum_{1\leq i<j\leq n} \overrightarrow{a_i}\bullet\overrightarrow{a_j} = -\frac{n}{2}$.

Problem 8.7.3 Let $a_k > 0$. Use the CBS Inequality to shew that

$$\left(\sum_{k=1}^n a_k^2\right)\left(\sum_{k=1}^n \frac{1}{a_k^2}\right) \geq n^2.$$

Problem 8.7.4 Let $\overrightarrow{a} \in \mathbb{R}^n$ be a fixed vector. Demonstrate that

$$X = \{\overrightarrow{x} \in \mathbb{R}^n : \overrightarrow{a}\bullet\overrightarrow{x} = 0\}$$

is a subspace of $\mathbb{R}^n$.

Problem 8.7.5 Let $\overrightarrow{a_i} \in \mathbb{R}^n$, $1 \leq i \leq k$ ($k \leq n$) be k non-zero vectors such that $\overrightarrow{a_i}\bullet\overrightarrow{a_j} = 0$ for $i \neq j$. Prove that these k vectors are linearly independent.

Problem 8.7.6 Let $a_k \geq 0, 1 \leq k \leq n$ be arbitrary. Prove that

$$\left(\sum_{k=1}^n a_k\right)^2 \leq \frac{n(n+1)(2n+1)}{6}\sum_{k=1}^n \frac{a_k^2}{k^2}.$$

Appendix A

Answers and Hints

1.1.2

$$\begin{aligned} x \in X \setminus (X \setminus A) &\iff x \in X \wedge x \notin (X \setminus A) \\ &\iff x \in X \wedge x \in A \\ &\iff x \in X \cap A. \end{aligned}$$

1.1.3

$$\begin{aligned} X \setminus (A \cup B) &\iff x \in X \wedge (x \notin (A \cup B)) \\ &\iff x \in X \wedge (x \notin A \wedge x \notin B) \\ &\iff (x \in X \wedge x \notin A) \wedge (x \in X \wedge x \notin B) \\ &\iff x \in (X \setminus A) \wedge x \in (X \setminus B) \\ &\iff x \in (X \setminus A) \cap (X \setminus B). \end{aligned}$$

1.1.6 One possible solution is

$$A \cup B \cup C = A \cup (B \setminus A) \cup (C \setminus (A \cup B)).$$

1.1.8 We have

$$|a| = |a - b + b| \leq |a - b| + |b|,$$

giving

$$|a| - |b| \leq |a - b|.$$

Similarly,

$$|b| = |b - a + a| \leq |b - a| + |a| = |a - b| + |a|,$$

gives

$$|b| - |a| \leq |a - b|.$$

The stated inequality follows from this.

1.2.1 $a \sim a$ since $\frac{a}{a} = 1 \in \mathbb{Z}$, and so the relation is reflexive. The relation is not symmetric. For $2 \sim 1$ since $\frac{2}{1} \in \mathbb{Z}$ but $1 \nsim 2$ since $\frac{1}{2} \notin \mathbb{Z}$. The relation is transitive. For assume $a \sim b$ and $b \sim c$. Then there exist $(m, n) \in \mathbb{Z}^2$ such that $\frac{a}{b} = m$, $\frac{b}{c} = n$. This gives

$$\frac{a}{c} = \frac{a}{b} \cdot \frac{b}{c} = mn \in \mathbb{Z},$$

and so $a \sim c$.

1.2.2 Here is one possible example: put $a \sim b \Leftrightarrow \frac{a^2+a}{b} \in \mathbb{Z}$. Then clearly if $a \in \mathbb{Z} \setminus \{0\}$ we have $a \sim a$ since $\frac{a^2+a}{a} = a + 1 \in \mathbb{Z}$. On the other hand, the relation is not symmetric, since $5 \sim 2$ as $\frac{5^2+5}{2} = 15 \in \mathbb{Z}$ but $2 \not\sim 5$, as $\frac{2^2+2}{5} = \frac{6}{5} \notin \mathbb{Z}$. It is not transitive either, since $\frac{5^2+5}{3} \in \mathbb{Z} \implies 5 \sim 3$ and $\frac{3^2+3}{12} \in \mathbb{Z} \implies 3 \sim 12$ but $\frac{5^2+5}{12} \notin \mathbb{Z}$ and so $5 \nsim 12$.

1.2.4 [B] $[x] = x + \frac{1}{3}\mathbb{Z}$. [C] No.

1.3.1 Let $\omega = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$. Then $\omega^2 + \omega + 1 = 0$ and $\omega^3 = 1$. Then

$$x = a^3 + b^3 + c^3 - 3abc = (a + b + c)(a + \omega b + \omega^2 c)(a + \omega^2 b + c\omega),$$

$$y = u^3 + v^3 + w^3 - 3uvw = (u + v + w)(u + \omega v + \omega^2 w)(u + \omega^2 v + \omega w).$$

Then

$$(a + b + c)(u + v + w) = au + av + aw + bu + bv + bw + cu + cv + cw,$$

$$\begin{aligned}(a + \omega b + \omega^2 c)(u + \omega v + \omega^2 w) &= au + bw + cv \\ &\quad + \omega(av + bu + cw) \\ &\quad + \omega^2(aw + bv + cu),\end{aligned}$$

and

$$\begin{aligned}(a + \omega^2 b + \omega c)(u + \omega^2 v + \omega w) &= au + bw + cv \\ &\quad + \omega(aw + bv + cu) \\ &\quad + \omega^2(av + bu + cw).\end{aligned}$$

This proves that

$$\begin{aligned}xy &= (au + bw + cv)^3 + (aw + bv + cu)^3 + (av + bu + cw)^3 \\ &\quad -3(au + bw + cv)(aw + bv + cu)(av + bu + cw),\end{aligned}$$

which proves that S is closed under multiplication.

1.3.2 We have

$$x\top(y\top z) = x\top(y \otimes a \otimes z) = (x) \otimes (a) \otimes (y \otimes a \otimes z) = x \otimes a \otimes y \otimes a \otimes z,$$

where we may drop the parentheses since $\otimes$ is associative. Similarly

$$(x\top y)\top z = (x \otimes a \otimes y)\top z = (x \otimes a \otimes y) \otimes (a) \otimes (z) = x \otimes a \otimes y \otimes a \otimes z.$$

By virtue of having proved

$$x\top(y\top z) = (x\top y)\top z,$$

associativity is established.

1.3.3 We proceed in order.

❶ Clearly, if a, b are rational numbers,

$$|a| < 1, |b| < 1 \implies |ab| < 1 \implies -1 < ab < 1 \implies 1 + ab > 0,$$

whence the denominator never vanishes and since sums, multiplications and divisions of rational numbers are rational, $\frac{a+b}{1+ab}$ is also rational. We must prove now that $-1 < \frac{a+b}{1+ab} < 1$ for $(a, b) \in]-1; 1[^2$. We have

$$\begin{aligned}-1 < \frac{a+b}{1+ab} < 1 &\Leftrightarrow -1 - ab < a + b < 1 + ab \\ &\Leftrightarrow -1 - ab - a - b < 0 < 1 + ab - a - b \\ &\Leftrightarrow -(a+1)(b+1) < 0 < (a-1)(b-1).\end{aligned}$$

Since $(a, b) \in]-1; 1[^2$, $(a+1)(b+1) > 0$ and so $-(a+1)(b+1) < 0$ giving the sinistral inequality. Similarly $a - 1 < 0$ and $b - 1 < 0$ give $(a-1)(b-1) > 0$, the dextral inequality. Since the steps are reversible, we have established that indeed $-1 < \frac{a+b}{1+ab} < 1$.

❷ Since $a \otimes b = \frac{a+b}{1+ab} = \frac{b+a}{1+ba} = b \otimes a$, commutativity follows trivially. Now

$$\begin{aligned} a \otimes (b \otimes c) &= a \otimes \left(\frac{b+c}{1+bc}\right) \\ &= \frac{a + \left(\frac{b+c}{1+bc}\right)}{1 + a\left(\frac{b+c}{1+bc}\right)} \\ &= \frac{a(1+bc)+b+c}{1+bc+a(b+c)} = \frac{a+b+c+abc}{1+ab+bc+ca}. \end{aligned}$$

One the other hand,

$$\begin{aligned} (a \otimes b) \otimes c &= \left(\frac{a+b}{1+ab}\right) \otimes c \\ &= \frac{\left(\frac{a+b}{1+ab}\right) + c}{1 + \left(\frac{a+b}{1+ab}\right)c} \\ &= \frac{(a+b)+c(1+ab)}{1+ab+(a+b)c} \\ &= \frac{a+b+c+abc}{1+ab+bc+ca}, \end{aligned}$$

whence $\otimes$ is associative.

❸ If $a \otimes e = a$ then $\frac{a+e}{1+ae} = a$, which gives $a + e = a + ea^2$ or $e(a^2 - 1) = 0$. Since $a \neq \pm 1$, we must have $e = 0$.

❹ If $a \otimes b = 0$, then $\frac{a+b}{1+ab} = 0$, which means that $b = -a$.

1.3.4 We proceed in order.

❶ Since $a \otimes b = a + b - ab = b + a - ba = b \otimes a$, commutativity follows trivially. Now

$$\begin{aligned} a \otimes (b \otimes c) &= a \otimes (b + c - bc) \\ &= a + b + c - bc - a(b + c - bc) \\ &= a + b + c - ab - bc - ca + abc. \end{aligned}$$

One the other hand,

$$\begin{aligned} (a \otimes b) \otimes c &= (a + b - ab) \otimes c \\ &= a + b - ab + c - (a + b - ab)c \\ &= a + b + c - ab - bc - ca + abc, \end{aligned}$$

whence $\otimes$ is associative.

❷ If $a \otimes e = a$ then $a + e - ae = a$, which gives $e(1 - a) = 0$. Since $a \neq 1$, we must have $e = 0$.

❸ If $a \otimes b = 0$, then $a + b - ab = 0$, which means that $b(1 - a) = -a$. Since $a \neq 1$ we find $b = -\frac{a}{1-a}$.

$+$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$
$\overline{0}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$
$\overline{1}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$
$\overline{2}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$
$\overline{3}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
$\overline{4}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$
$\overline{5}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$
$\overline{6}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$
$\overline{7}$	$\overline{7}$	$\overline{0}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$
$\overline{8}$	$\overline{8}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$
$\overline{9}$	$\overline{9}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$
$\overline{10}$	$\overline{10}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$

Table A.1: Addition table for $\mathbb{Z}_{11}$.

$\cdot$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$
$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$	$\overline{0}$
$\overline{1}$	$\overline{0}$	$\overline{1}$	$\overline{2}$	$\overline{3}$	$\overline{4}$	$\overline{5}$	$\overline{6}$	$\overline{7}$	$\overline{8}$	$\overline{9}$	$\overline{10}$
$\overline{2}$	$\overline{0}$	$\overline{2}$	$\overline{4}$	$\overline{6}$	$\overline{8}$	$\overline{10}$	$\overline{1}$	$\overline{3}$	$\overline{5}$	$\overline{7}$	$\overline{9}$
$\overline{3}$	$\overline{0}$	$\overline{3}$	$\overline{6}$	$\overline{9}$	$\overline{1}$	$\overline{4}$	$\overline{7}$	$\overline{10}$	$\overline{2}$	$\overline{5}$	$\overline{8}$
$\overline{4}$	$\overline{0}$	$\overline{4}$	$\overline{8}$	$\overline{1}$	$\overline{5}$	$\overline{9}$	$\overline{2}$	$\overline{6}$	$\overline{10}$	$\overline{3}$	$\overline{7}$
$\overline{5}$	$\overline{0}$	$\overline{5}$	$\overline{10}$	$\overline{4}$	$\overline{9}$	$\overline{3}$	$\overline{8}$	$\overline{2}$	$\overline{7}$	$\overline{1}$	$\overline{6}$
$\overline{6}$	$\overline{0}$	$\overline{6}$	$\overline{1}$	$\overline{7}$	$\overline{2}$	$\overline{8}$	$\overline{3}$	$\overline{9}$	$\overline{4}$	$\overline{10}$	$\overline{5}$
$\overline{7}$	$\overline{0}$	$\overline{7}$	$\overline{3}$	$\overline{10}$	$\overline{6}$	$\overline{2}$	$\overline{9}$	$\overline{5}$	$\overline{1}$	$\overline{8}$	$\overline{4}$
$\overline{8}$	$\overline{0}$	$\overline{8}$	$\overline{5}$	$\overline{2}$	$\overline{10}$	$\overline{7}$	$\overline{4}$	$\overline{1}$	$\overline{9}$	$\overline{6}$	$\overline{3}$
$\overline{9}$	$\overline{0}$	$\overline{9}$	$\overline{7}$	$\overline{5}$	$\overline{3}$	$\overline{1}$	$\overline{10}$	$\overline{8}$	$\overline{6}$	$\overline{4}$	$\overline{2}$
$\overline{10}$	$\overline{0}$	$\overline{10}$	$\overline{9}$	$\overline{8}$	$\overline{7}$	$\overline{6}$	$\overline{5}$	$\overline{4}$	$\overline{3}$	$\overline{2}$	$\overline{1}$

Table A.2: Multiplication table $\mathbb{Z}_{11}$.

1.3.5 We have

$$\begin{aligned} x \circ y &= (x \circ y) \circ (x \circ y) \\ &= [y \circ (x \circ y)] \circ x \\ &= [(x \circ y) \circ x] \circ y \\ &= [(y \circ x) \circ x] \circ y \\ &= [(x \circ x) \circ y] \circ y \\ &= (y \circ y) \circ (x \circ x) \\ &= y \circ x, \end{aligned}$$

proving commutativity.

1.4.1 The tables appear in tables A.1 and A.2.

1.4.2 Observe that

$$\overline{3}x^2 - \overline{5}x + \overline{1} = \overline{0} \implies \overline{4}(\overline{3}x^2 - \overline{5}x + \overline{1}) = \overline{40} \implies x^2 + \overline{2}x + \overline{1} + \overline{3} = \overline{0} \implies (x + \overline{1})^2 = \overline{8}.$$

We need to know whether $\overline{8}$ is a perfect square modulo 11. Observe that $(\overline{11 - a})^2 = \overline{a}$, so we just need to check half the elements and see that

$$\overline{1}^2 = \overline{1}; \quad \overline{2}^2 = \overline{4}; \quad \overline{3}^2 = \overline{9}; \quad \overline{4}^2 = \overline{5}; \quad \overline{5}^2 = \overline{3},$$

whence $\overline{8}$ is not a perfect square modulo 11 and so there are no solutions.

1.4.3 From example 50

$$x^2 = \overline{5}.$$

Now, the squares modulo 11 are $\overline{0}^2 = \overline{0}$, $\overline{1}^2 = \overline{1}$, $\overline{2}^2 = \overline{4}$, $\overline{3}^2 = \overline{9}$, $\overline{4}^2 = \overline{5}$, $\overline{5}^2 = \overline{3}$. Also, $(\overline{11 - 4})^2 = \overline{7}^2 = \overline{5}$. Hence the solutions are $x = \overline{4}$ or $x = \overline{7}$.

1.4.5 Put $f(x) = x^4 + x^3 + x^2 + x + \overline{1}$. Then

$f(0) = 1 \equiv 1 \mod 11$	$f(1) = 5 \equiv 5 \mod 11$	$f(2) = 31 \equiv 9 \mod 11$
$f(3) = 121 \equiv 0 \mod 11$	$f(4) = 341 \equiv 0 \mod 11$	$f(5) = 781 \equiv 0 \mod 11$
$f(6) = 1555 \equiv 4 \mod 11$	$f(7) = 2801 \equiv 7 \mod 11$	$f(8) = 4681 \equiv 6 \mod 11$
$f(9) = 7381 \equiv 0 \mod 11$	$f(10) = 11111 \equiv 1 \mod 11$	

1.5.1 We have

$$\begin{aligned}\frac{1}{\sqrt{2}+2\sqrt{3}+3\sqrt{6}} &= \frac{\sqrt{2}+2\sqrt{3}-3\sqrt{6}}{(\sqrt{2}+2\sqrt{3})^2-(3\sqrt{6})^2}\\ &= \frac{\sqrt{2}+2\sqrt{3}-3\sqrt{6}}{2+12+4\sqrt{6}-54}\\ &= \frac{\sqrt{2}+2\sqrt{3}-3\sqrt{6}}{-40+4\sqrt{6}}\\ &= \frac{(\sqrt{2}+2\sqrt{3}-3\sqrt{6})(-40-4\sqrt{6})}{40^2-(4\sqrt{6})^2}\\ &= \frac{(\sqrt{2}+2\sqrt{3}-3\sqrt{6})(-40-4\sqrt{6})}{1504}\\ &= -\frac{16\sqrt{2}+22\sqrt{3}-30\sqrt{6}-18}{376}\end{aligned}$$

1.5.2 Since

$$(-a)b^{-1}+ab^{-1}=(-a+a)b^{-1}=0_{\mathbb{F}}b^{-1}=0_{\mathbb{F}},$$

we obtain by adding $-(ab^{-1})$ to both sides that

$$(-a)b^{-1}=-(ab^{-1}).$$

Similarly, from

$$a(-b^{-1})+ab^{-1}=a(-b^{-1}+b^{-1})=a0_{\mathbb{F}}=0_{\mathbb{F}},$$

we obtain by adding $-(ab^{-1})$ to both sides that

$$a(-b^{-1})=-(ab^{-1}).$$

1.6.1 Assume $h(b)=h(a)$. Then

$$\begin{aligned}h(a)=h(b) &\implies a^3=b^3\\ &\implies a^3-b^3=0\\ &\implies (a-b)(a^2+ab+b^2)=0\end{aligned}$$

Now,

$$b^2+ab+a^2=\left(b+\frac{a}{2}\right)^2+\frac{3a^2}{4}.$$

This shews that b^2+ab+a^2 is positive unless both a and b are zero. Hence $a-b=0$ in all cases. We have shewn that $h(b)=h(a) \implies a=b$, and the function is thus injective.

1.6.2 We have

$$\begin{aligned}f(a)=f(b) &\iff \frac{6a}{2a-3}=\frac{6b}{2b-3}\\ &\iff 6a(2b-3)=6b(2a-3)\\ &\iff 12ab-18a=12ab-18b\\ &\iff -18a=-18b\\ &\iff a=b,\end{aligned}$$

proving that f is injective. Now, if

$$f(x)=y,\quad y\neq 3,$$

then

$$\frac{6x}{2x-3}=y,$$

that is $6x=y(2x-3)$. Solving for x we find

$$x=\frac{3y}{2y-6}.$$

Since $2y-6\neq 0$, x is a real number, and so f is surjective. On combining the results we deduce that f is bijective.

2.1.1 $A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 4 & 8 \\ 3 & 9 & 27 \end{bmatrix}$.

2.1.2 $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$.

2.1.3 $M + N = \begin{bmatrix} a+1 & 0 & 2c \\ a & b-2a & 0 \\ 2a & 0 & -2 \end{bmatrix}$, $2M = \begin{bmatrix} 2a & -4a & 2c \\ 0 & -2a & 2b \\ 2a+2b & 0 & -2 \end{bmatrix}$.

2.1.4 $x = 1$ and $y = 4$.

2.1.5 $A = \begin{bmatrix} 13 & -1 \\ 15 & 3 \end{bmatrix}$, $B = \begin{bmatrix} 5 & 0 \\ 6 & 1 \end{bmatrix}$

2.1.8 The set of border elements is the union of two rows and two columns. Thus we may choose at most four elements from the border, and at least one from the central 3×3 matrix. The largest element of this 3×3 matrix is 15, so any allowable choice of does not exceed 15. The choice 25, 15, 18,l 23, 20 shews that the largest minimum is indeed 15.

2.2.1 $\begin{bmatrix} 2 & 2 \\ 0 & -2 \end{bmatrix}$

2.2.2

$$AB = \begin{bmatrix} a & b & c \\ c+a & a+b & b+c \\ a+b+c & a+b+c & a+b+c \end{bmatrix}, \quad BA = \begin{bmatrix} a+b+c & b+c & c \\ a+b+c & a+b & b \\ a+b+c & c+a & a \end{bmatrix}$$

2.2.3 $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 14 & 14 & 14 \\ 11 & 11 & 11 \\ 11 & 11 & 11 \end{bmatrix}$, whence $a + b + c = 36$.

2.2.4 An easy computation leads to $N^2 = \begin{bmatrix} 0 & 0 & 4 & 12 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $N^3 = \begin{bmatrix} 0 & 0 & 0 & -8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ and $N^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Hence any power—from the fourth on—is the zero matrix.

2.2.5 $AB = \mathbf{0}_4$ and $BA = \begin{bmatrix} \overline{0} & \overline{2} & \overline{0} & \overline{3} \\ \overline{0} & \overline{2} & \overline{0} & \overline{3} \\ \overline{0} & \overline{2} & \overline{0} & \overline{3} \\ \overline{0} & \overline{2} & \overline{0} & \overline{3} \end{bmatrix}$.

2.2.6 For the first part, observe that

$$\begin{aligned} m(a)m(b) &= \begin{bmatrix} 1 & 0 & a \\ -a & 1 & -\frac{a^2}{2} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & b \\ -b & 1 & -\frac{b^2}{2} \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & a+b \\ -a-b & 1 & -\frac{a^2}{2} - \frac{b^2}{2} + ab \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & a+b \\ -(a+b) & 1 & -\frac{(a+b)^2}{2} \\ 0 & 0 & 1 \end{bmatrix} \\ &= m(a+b) \end{aligned}$$

For the second part, observe that using the preceding part of the problem,

$$m(a)m(-a) = m(a-a) = m(0) = \begin{bmatrix} 1 & 0 & 0 \\ -0 & 1 & -\frac{0^2}{2} \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{I}_3,$$

giving the result.

2.2.7 Observe that

$$A^2 = (AB)(AB) = A(BA)B = A(B)B = (AB)B = AB = A.$$

Similarly,

$$B^2 = (BA)(BA) = B(AB)A = B(A)A = (BA)A = BA = B.$$

2.2.8 For this problem you need to recall that if $|r| < 1$, then

$$a + ar + ar^2 + ar^3 + \cdots + \cdots = \frac{a}{1-r}.$$

This gives

$$1 + \tfrac{1}{4} + \tfrac{1}{4^2} + \tfrac{1}{4^3} + \cdots = \frac{1}{1-\frac{1}{4}} = \frac{4}{3},$$

$$\tfrac{1}{2} + \tfrac{1}{2^3} + \tfrac{1}{2^5} + \cdots = \frac{\frac{1}{2}}{1-\frac{1}{4}} = \frac{2}{3},$$

and

$$1 + \tfrac{1}{2} + \tfrac{1}{2^2} + \tfrac{1}{2^3} + \cdots = \frac{1}{1-\frac{1}{2}} = 2.$$

By looking at a few small cases, it is easy to establish by induction that for $n \geq 1$

$$A^{2n-1} = \begin{bmatrix} 0 & \frac{1}{2^{2n-1}} & 0 \\ \frac{1}{2^{2n-1}} & 0 & 0 \\ 0 & 0 & \frac{1}{2^{2n-1}} \end{bmatrix}, \qquad A^{2n} = \begin{bmatrix} \frac{1}{2^{2n}} & 0 & 0 \\ 0 & \frac{1}{2^{2n}} & 0 \\ 0 & 0 & \frac{1}{2^{2n}} \end{bmatrix}.$$

This gives

$$\mathbf{I}_3 + A + A^2 + A^3 + \cdots = \begin{bmatrix} 1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \cdots & \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \cdots & 0 \\ \frac{1}{2} + \frac{1}{2^3} + \frac{1}{2^5} + \cdots & 1 + \frac{1}{4} + \frac{1}{4^2} + \frac{1}{4^3} + \cdots & 0 \\ 0 & 0 & 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots \end{bmatrix} = \begin{bmatrix} \frac{4}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{4}{3} & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

2.2.9 Observe that

$$\begin{bmatrix} -4 & x \\ -x & 4 \end{bmatrix}^2 = \begin{bmatrix} -4 & x \\ -x & 4 \end{bmatrix}\begin{bmatrix} -4 & x \\ -x & 4 \end{bmatrix} = \begin{bmatrix} 16 - x^2 & 0 \\ 0 & 16 - x^2 \end{bmatrix},$$

and so we must have $16 - x^2 = -1$ or $x = \pm\sqrt{17}$.

2.2.10 Disprove! Take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $AB = B$, but $BA = \mathbf{0}_2$.

2.2.11 Disprove! Take for example $A = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$. Then

$$A^2 - B^2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \neq \begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix} = (A + B)(A - B).$$

2.2.12 $x = 6$.

2.2.14 $\begin{bmatrix} 32 & -32 \\ -32 & 32 \end{bmatrix}$.

2.2.15 $A^{2003} = \begin{bmatrix} 0 & 2^{1001}3^{1002} \\ 2^{1002}3^{1001} & 0 \end{bmatrix}$.

2.2.17 The assertion is clearly true for $n = 1$. Assume that is it true for n, that is, assume

$$A^n = \begin{bmatrix} \mathbf{cos}(n)\alpha & -\mathbf{sin}(n)\alpha \\ \mathbf{sin}(n)\alpha & \mathbf{cos}(n)\alpha \end{bmatrix}.$$

Then

$$\begin{aligned} A^{n+1} &= AA^n \\ &= \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \begin{bmatrix} \cos(n)\alpha & -\sin(n)\alpha \\ \sin(n)\alpha & \cos(n)\alpha \end{bmatrix} \\ &= \begin{bmatrix} \cos\alpha\cos(n)\alpha - \sin\alpha\sin(n)\alpha & -\cos\alpha\sin(n)\alpha - \sin\alpha\cos(n)\alpha \\ \sin\alpha\cos(n)\alpha + \cos\alpha\sin(n)\alpha & -\sin\alpha\sin(n)\alpha + \cos\alpha\cos(n)\alpha \end{bmatrix} \\ &= \begin{bmatrix} \cos(n+1)\alpha & -\sin(n+1)\alpha \\ \sin(n+1)\alpha & \cos(n+1)\alpha \end{bmatrix}, \end{aligned}$$

and the result follows by induction.

2.2.18 Let $A = [a_{ij}], B = [b_{ij}]$ be checkered $n \times n$ matrices. Then $A + B = (a_{ij} + b_{ij})$. If $j - i$ is odd, then $a_{ij} + b_{ij} = 0 + 0 = 0$, which shows that $A + B$ is checkered. Furthermore, let $AB = [c_{ij}]$ with $c_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. If i is even and j odd, then $a_{ik} = 0$ for odd k and $b_{kj} = 0$ for even k. Thus $c_{ij} = 0$ for i even and j odd. Similarly, $c_{ij} = 0$ for odd i and even j. This proves that AB is checkered.

2.2.19 Put

$$J = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

We first notice that

$$J^2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad J^3 = \mathbf{0}_3.$$

This means that the sum in the binomial expansion

$$A^n = (\mathbf{I}_3 + J)^n = \sum_{k=0}^{n} \binom{n}{k} \mathbf{I}^{n-k} J^k$$

is a sum of zero matrices for $k \geq 3$. We thus have

$$\begin{aligned} A^n &= \mathbf{I}_3^n + n\mathbf{I}_3^{n-1}J + \tbinom{n}{2}\mathbf{I}_3^{n-2}J^2 \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & n & n \\ 0 & 0 & n \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \binom{n}{2} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & n & \frac{n(n+1)}{2} \\ 0 & 1 & n \\ 0 & 0 & 1 \end{bmatrix}, \end{aligned}$$

giving the result, since $\binom{n}{2} = \frac{n(n-1)}{2}$ and $n + \binom{n}{2} = \frac{n(n+1)}{2}$.

2.2.20 Argue inductively,

$$A^2B = A(AB) = AB = B$$

$$A^3B = A(A^2B) = A(AB) = AB = B$$

$$\vdots$$

$$A^mB = AB = B.$$

Hence $B = A^mB = \mathbf{0}_nB = \mathbf{0}_n$.

2.2.22 Put $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Using 2.2.21, deduce by iteration that

$$A^k = (a+d)^{k-1}A.$$

2.2.23 $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, $bc = -a^2$

2.2.24 $\pm I_2$, $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$, $a^2 = 1 - bc$

2.2.25 We complete squares by putting $Y = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = X - I$. Then

$$\begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2 \end{bmatrix} = Y^2 = X^2 - 2X + I = (X-I)^2 = \begin{bmatrix} -1 & 0 \\ 6 & 3 \end{bmatrix} + I = \begin{bmatrix} 0 & 0 \\ 6 & 4 \end{bmatrix}.$$

This entails $a = 0, b = 0, cd = 6, d^2 = 4$. Using $X = Y + I$, we find that there are two solutions,

$$\begin{bmatrix} 1 & 0 \\ 3 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ -3 & -1 \end{bmatrix}.$$

2.2.26 The matrix

$$A = \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

clearly satisfies the conditions.

2.2.27 Put $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then

$$X^2 + X = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \iff \begin{cases} a^2 + bc + a = 1 \\ ab + bd + b = 1 \\ ca + dc + c = 1 \\ cb + d^2 + d = 1 \end{cases} \iff \begin{cases} a^2 + bc + a = 1 \\ b(a + d + 1) = 1 \\ c(a + d + 1) = 1 \\ (d - a)(a + d + 1) = 0 \end{cases} \iff \begin{cases} d = a \neq \frac{1}{2} \\ c = b = \frac{1}{2a + 1} \\ a^2 + \frac{1}{(2a+1)^2} + a = 1 \end{cases}$$

The last equation holds

$$\iff 4a^4 + 8a^3 + a^2 - 3a = 0 \iff a \in \{-\frac{3}{2}, -1, 0, \frac{1}{2}\}.$$

Thus the set of solutions is

$$\{\begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \begin{bmatrix} -3/2 & -1/2 \\ -1/2 & -3/2 \end{bmatrix}\}$$

2.2.29 Observe that $\mathbf{A} = 2\mathbf{I}_3 - J$, where $\mathbf{I}_3$ is the 3×3 identity matrix and

$$J = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Observe that $J^k = 3^{k-1}J$ for integer $k \geq 1$. Using the binomial theorem we have

$$\begin{aligned} \mathbf{A}^n &= (2\mathbf{I}_3 - J)^n \\ &= \sum_{k=0}^{n} \binom{n}{k} (2\mathbf{I}_3)^{n-k}(-1)^k J^k \\ &= 2^n\mathbf{I}_3 + \frac{1}{3}J \sum_{k=1}^{n} \binom{n}{k} 2^{n-k}(-1)^k 3^k \\ &= 2^n\mathbf{I}_3 + \frac{1}{3}J((-1)^n - 2^n) \\ &= \frac{1}{3}\begin{bmatrix} (-1)^n + 2^{n+1} & (-1)^n - 2^n & (-1)^n - 2^n \\ (-1)^n - 2^n & (-1)^n + 2^{n+1} & (-1)^n - 2^n \\ (-1)^n - 2^n & (-1)^n - 2^n & (-1)^n + 2^{n+1} \end{bmatrix}. \end{aligned}$$

2.3.1 There are infinitely many solutions. Here is one:

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

2.3.2 There are infinitely many examples. One could take $\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. Another set is

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}.$$

2.3.3 If such matrices existed, then by the first equation

$$\mathbf{tr}(AC) + \mathbf{tr}(DB) = n.$$

By the second equation and by Theorem 86,

$$0 = \mathbf{tr}(CA) + \mathbf{tr}(BD) = \mathbf{tr}(AC) + \mathbf{tr}(DB) = n,$$

a contradiction, since $n \geq 1$.

2.3.4 Disprove! This is not generally true. Take $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$. Clearly $A^T = A$ and $B^T = B$. We have

$$AB = \begin{bmatrix} 3 & 1 \\ 3 & 2 \end{bmatrix}$$

but

$$(AB)^T = \begin{bmatrix} 3 & 3 \\ 1 & 2 \end{bmatrix}.$$

2.3.6 We have

$$\mathbf{tr}\left(A^2\right) = \mathbf{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \mathbf{tr}\left(\begin{bmatrix} a^2 + bc & ab + bd \\ ca + cd & d^2 + cb \end{bmatrix}\right) = a^2 + d^2 + 2bc$$

and

$$\left(\mathbf{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right)\right)^2 = (a + d)^2.$$

Thus

$$\mathbf{tr}\left(A^2\right) = (\mathbf{tr}(A))^2 \iff a^2 + d^2 + 2bc = (a + d)^2 \iff bc = ad,$$

is the condition sought.

2.3.7

$$\begin{aligned} \mathbf{tr}\left((A - \mathbf{I}_4)^2\right) &= \mathbf{tr}\left(A^2 - 2A + \mathbf{I}_4\right) \\ &= \mathbf{tr}\left(A^2\right) - 2\mathbf{tr}(A) + \mathbf{tr}(\mathbf{I}_4) \\ &= -4 - 2\mathbf{tr}(A) + 4 \\ &= -2\mathbf{tr}(A), \end{aligned}$$

and $\mathbf{tr}(3\mathbf{I}_4) = 12$. Hence $-2\mathbf{tr}(A) = 12$ or $\mathbf{tr}(A) = -6$.

2.3.8 Disprove! Take $A = B = \mathbf{I}_n$ and $n > 1$. Then $\mathbf{tr}(AB) = n < n^2 = \mathbf{tr}(A)\,\mathbf{tr}(B)$.

2.3.9 Disprove! Take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $C = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$. Then $\mathbf{tr}(ABC) = 1$ but $\mathbf{tr}(BAC) = 0$.

2.3.10 We have

$$(AA^T)^T = (A^T)^T A^T = AA^T.$$

2.3.11 We have

$$(AB - BA)^T = (AB)^T - (BA)^T = B^T A^T - A^T B^T = -BA - A(-B) = AB - BA.$$

2.3.13 Let $X = [x_{ij}]$ and put $XX^T = [c_{ij}]$. Then

$$0 = c_{ii} = \sum_{k=1}^{n} x_{ik}^2 \implies x_{ik} = 0.$$

2.4.1 Here is one possible approach. If we perform $C_1 \leftrightarrow C_3$ on A we obtain

$$A_1 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix} \quad \textbf{so take} \quad P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now perform $2R_1 \to R_1$ on A_1 to obtain

$$A_2 = \begin{bmatrix} 2 & 0 & 2 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix} \quad \textbf{so take} \quad D = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Finally, perform $R_1 + 2R_4 \to R_1$ on A_2 to obtain

$$B = \begin{bmatrix} 4 & -2 & 4 & 2 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \end{bmatrix} \quad \textbf{so take} \quad T = \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

2.4.2 Here is one possible approach.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \overset{P:\rho_3 \leftrightarrow \rho_1}{\rightsquigarrow} \begin{bmatrix} g & h & i \\ d & e & f \\ a & b & c \end{bmatrix}$$

$$\overset{P':C_1 \leftrightarrow C_2}{\rightsquigarrow} \begin{bmatrix} h & g & i \\ e & d & f \\ b & a & c \end{bmatrix}$$

$$\overset{T:C_1 - C_2 \to C_1}{\rightsquigarrow} \begin{bmatrix} h-g & g & i \\ e-d & d & f \\ b-a & a & c \end{bmatrix}$$

$$\overset{D:2\rho_3 \to \rho_3}{\rightsquigarrow} \begin{bmatrix} h-g & g & i \\ e-d & d & f \\ 2b-2a & 2a & 2c \end{bmatrix}$$

Thus we take

$$P = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad P' = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$T = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

2.4.3 Let $\mathbf{E}_{ij} \in \mathbf{M}_{n\times n}(\mathbb{F})$. Then

$$\mathbf{AE}_{ij} = \begin{bmatrix} 0 & 0 & \dots & a_{1i} & \dots & 0 \\ 0 & 0 & \vdots & a_{2i} & \vdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \vdots & a_{n-1i} & \vdots & 0 \\ 0 & 0 & \vdots & a_{ni} & \vdots & 0 \end{bmatrix},$$

where the entries appear on the j-column. Then we see that $\mathbf{tr}(\mathbf{AE}_{ij}) = a_{ji}$ and similarly, by considering $\mathbf{BE}_{ij}$, we see that $\mathbf{tr}(\mathbf{BE}_{ij}) = b_{ji}$. Therefore $\forall i, j, \quad a_{ji} = b_{ji}$, which implies that $A = B$.

2.4.4 Let $\mathbf{E}_{st} \in \mathbf{M}_{n\times n}(\mathbb{R})$. Then

$$\mathbf{E}_{ij}A = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ a_{j1} & a_{j2} & \dots & a_{jn} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix},$$

where the entries appear on the i-th row. Thus

$$(\mathbf{E}_{ij}A)^2 = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ a_{ji}a_{j1} & a_{ji}a_{j2} & \dots & a_{ji}a_{jn} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix},$$

which means that $\forall i, j, a_{ji}a_{jk} = 0$. In particular, $a_{ji}^2 = 0$, which means that $\forall i, j, a_{ji} = 0$, i.e., $A = \mathbf{0}_n$.

2.5.1 $a = 1$, $b = -2$.

2.5.2 Claim: $A^{-1} = \mathbf{I}_n - A + A^2 - A^3$. For observe that

$$(\mathbf{I}_n + A)(\mathbf{I}_n - A + A^2 - A^3) = \mathbf{I}_n - A + A^2 - A^3 - A + A^2 - A^3 + A^4 = \mathbf{I}_n,$$

proving the claim.

2.5.3 Disprove! It is enough to take $A = B = 2\mathbf{I}_n$. Then $(A+B)^{-1} = (4\mathbf{I}_n)^{-1} = \frac{1}{4}\mathbf{I}_n$ but $A^{-1} + B^{-1} = \frac{1}{2}\mathbf{I}_n + \frac{1}{2}\mathbf{I}_n = \mathbf{I}_n$.

2.5.8 We argue by contradiction. If exactly one of the matrices is not invertible, the identities

$$A = A\mathbf{I}_n = (ABC)(BC)^{-1} = \mathbf{0}_n,$$

$$B = \mathbf{I}_n B\mathbf{I}_n = (A)^{-1}(ABC)C^{-1} = \mathbf{0}_n,$$

$$C = \mathbf{I}_n C = (AB)^{-1}(ABC) = \mathbf{0}_n,$$

shew a contradiction depending on which of the matrices are invertible. If all the matrices are invertible then

$$\mathbf{0}_n = \mathbf{0}_n C^{-1}B^{-1}A^{-1} = (ABC)C^{-1}B^{-1}A^{-1} = \mathbf{I}_n,$$

also gives a contradiction.

2.5.9 Observe that A, B, AB are invertible. Hence

$$A^2B^2 = \mathbf{I}_n = (AB)^2 \implies AABB = ABAB$$

$$\implies AB = BA,$$

by cancelling A on the left and B on the right. One can also argue that $A = A^{-1}$, $B = B^{-1}$, and so $AB = (AB)^{-1} = B^{-1}A^{-1} = BA$.

2.5.10 Observe that $A = (a-b)\mathbf{I}_n + bU$, where U is the $n \times n$ matrix with $1_{\mathbb{F}}$'s everywhere. Prove that

$$A^2 = (2(a-b) + nb)A - ((a-b)^2 + nb(a-b))\mathbf{I}_n.$$

2.5.11 Compute $(A - \mathbf{I}_n)(B - \mathbf{I}_n)$.

2.5.12 By Theorem 86 we have $\mathbf{tr}\left(SAS^{-1}\right) = \mathbf{tr}\left(S^{-1}SA\right) = \mathbf{tr}\left(A\right)$.

2.7.2 The rank is 2.

2.7.3 If B is invertible, then $\mathbf{rank}\left(AB\right) = \mathbf{rank}\left(A\right) = \mathbf{rank}\left(BA\right)$. Similarly, if A is invertible $\mathbf{rank}\left(AB\right) = \mathbf{rank}\left(B\right) = \mathbf{rank}\left(BA\right)$. Now, take $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Then $AB = B$, and so $\mathbf{rank}\left(AB\right) = 1$. But $BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, and so $\mathbf{rank}\left(BA\right) = 0$.

2.7.4 Observe that

$$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 2 & 0 & 0 & 2 \end{bmatrix} \overset{R_3-2(R_1+R_2)\to R_3}{\underset{R_4-2R_1\to R_4}{\rightsquigarrow}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & -2 & 0 & 2 \end{bmatrix},$$

whence the matrix has three pivots and so rank 3.

2.7.5 The maximum rank of this matrix could be 2. Hence, for the rank to be 1, the rows must be proportional, which entails

$$\frac{x^2}{4} = \frac{x}{2} \implies x^2 - 2x = 0 \implies x \in \{0, 2\}.$$

2.7.6 Assume first that the non-zero $n \times n$ matrix A over a field $\mathbb{F}$ has rank 1. By permuting the rows of the matrix we may assume that every other row is a scalar multiple of the first row, which is non-zero since the rank is 1. Hence A must be of the form

$$A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \\ \lambda_1 a_1 & \lambda_1 a_2 & \cdots & \lambda_1 a_n \\ \vdots & \vdots & \cdots & \vdots \\ \lambda_{n-1} a_1 & \lambda_{n-1} a_2 & \cdots & \lambda_{n-1} a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_{n-1} \end{bmatrix} := XY,$$

which means that the claimed factorisation indeed exists.

Conversely, assume that A can be factored as $A = XY$, where $X \in \mathbf{M}_{n\times 1}(\mathbb{F})$ and $Y \in \mathbf{M}_{1\times n}(\mathbb{F})$. Since A is non-zero, we must have $\mathbf{rank}\left(A\right) \geq 1$. Similarly, neither X nor Y could be all zeroes, because otherwise A would be zero. This means that $\mathbf{rank}\left(X\right) = 1 = \mathbf{rank}\left(Y\right)$. Now, since

$$\mathbf{rank}\left(A\right) \leq \mathbf{min}(\mathbf{rank}\left(X\right), \mathbf{rank}\left(Y\right)) = 1,$$

we deduce that $\mathbf{rank}\left(A\right) \leq 1$, proving that $\mathbf{rank}\left(A\right) = 1$.

2.7.7 Effecting $R_3 - R_1 \to R_3$; $aR_4 - bR_2 \to R_4$ successively, we obtain

$$\begin{bmatrix} 1 & a & 1 & b \\ a & 1 & b & 1 \\ 1 & b & 1 & a \\ b & 1 & a & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & a & 1 & b \\ a & 1 & b & 1 \\ 0 & b-a & 0 & a-b \\ 0 & a-b & a^2-b^2 & a-b \end{bmatrix}.$$

Performing $R_2 - aR_1 \to R_2$; $R_4 + R_3 \to R_4$ we have

$$\rightsquigarrow \begin{bmatrix} 1 & a & 1 & b \\ 0 & 1-a^2 & b-a & 1-ab \\ 0 & b-a & 0 & a-b \\ 0 & 0 & a^2-b^2 & 2(a-b) \end{bmatrix}.$$

Performing $(1-a^2)R_3 - (b-a)R_2 \to R_3$ we have

$$\rightsquigarrow \begin{bmatrix} 1 & a & 1 & b \\ 0 & 1-a^2 & b-a & 1-ab \\ 0 & 0 & -a^2+2ab-b^2 & 2a-2b-a^3+ab^2 \\ 0 & 0 & a^2-b^2 & 2(a-b) \end{bmatrix}.$$

Performing $R_3 - R_4 \to R_3$ we have

$$\rightsquigarrow \begin{bmatrix} 1 & a & 1 & b \\ 0 & 1-a^2 & b-a & 1-ab \\ 0 & 0 & -2a(a-b) & -a(a^2-b^2) \\ 0 & 0 & a^2-b^2 & 2(a-b) \end{bmatrix}.$$

Performing $2aR_4 + (a+b)R_3 \to R_4$ we have

$$\begin{bmatrix} 1 & a & 1 & b \\ 0 & 1-a^2 & b-a & 1-ab \\ 0 & 0 & -2a(a-b) & -a(a^2-b^2) \\ 0 & 0 & 0 & 4a^2-4ab-a^4+a^2b^2-ba^3+ab^3 \end{bmatrix}.$$

Factorising, this is

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

Thus the rank is 4 if $(a+2+b)(a-b)(a-2+b) \neq 0$. The rank is 3 if $a+b=2$ and $(a,b) \neq (1,1)$ or if $a+b=-2$ and $(a,b) \neq (-1,-1)$. The rank is 2 if $a=b\neq 1$ and $a \neq -1$. The rank is 1 if $a=b=\pm 1$.

2.7.8 **rank** $(A) = 4$ if $m^3+m^2+2 \neq 0$, and **rank** $(A) = 3$ otherwise.

2.7.9 The rank is 4 if $a \neq \pm b$. The rank is 1 is $a = \pm b \neq 0$. The rank is 0 if $a=b=0$.

2.7.10 The rank is 4 if $(a-b)(c-d) \neq 0$. The rank is 2 is $a=b$, $c \neq d$ or if $a \neq b$, $c=d$. The rank is 1 if $a=b$ and $c=d$.

2.7.11 Observe that $\textbf{rank}(ABC) \leq \textbf{rank}(B) \leq 2$. Now,

$$\begin{bmatrix} 1 & 1 & 2 \\ -2 & x & 1 \\ 1 & -2 & 1 \end{bmatrix} \underset{R_3 - R_1 \to R_3}{\overset{R_2 + 2R_1 \to R_2}{\rightsquigarrow}} \begin{bmatrix} 1 & 1 & 2 \\ 0 & x+2 & 5 \\ 0 & -3 & -1 \end{bmatrix},$$

has rank at least 2, since the first and third rows are not proportional. This means that it must have rank exactly two, and the last two rows must be proportional. Hence

$$\frac{x+2}{-3} = \frac{5}{-1} \implies x = 13.$$

2.7.13 For the counterexample consider $A = \begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$.

2.8.1 We form the augmented matrix

$$\left[\begin{array}{ccc|ccc} \overline{1} & \overline{2} & \overline{3} & \overline{1} & \overline{0} & \overline{0} \\ \overline{2} & \overline{3} & \overline{1} & \overline{0} & \overline{1} & \overline{0} \\ \overline{3} & \overline{1} & \overline{2} & \overline{0} & \overline{0} & \overline{1} \end{array}\right]$$

From $R_2 - \overline{2}R_1 \to R_2$ and $R_3 - \overline{3}R_1 \to R_3$ we obtain

$$\rightsquigarrow \left[\begin{array}{ccc|ccc} \overline{1} & \overline{2} & \overline{3} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{6} & \overline{2} & \overline{5} & \overline{1} & \overline{0} \\ \overline{0} & \overline{2} & \overline{0} & \overline{4} & \overline{0} & \overline{1} \end{array}\right].$$

From $R_2 \leftrightarrow R_3$ we obtain

$$\rightsquigarrow \left[\begin{array}{ccc|ccc} \overline{1} & \overline{2} & \overline{3} & \overline{1} & \overline{0} & \overline{0} \\ \overline{0} & \overline{2} & \overline{0} & \overline{4} & \overline{0} & \overline{1} \\ \overline{0} & \overline{6} & \overline{2} & \overline{5} & \overline{1} & \overline{0} \end{array}\right].$$

Now, from $R_1 - R_2 \to R_1$ and $R_3 - \overline{3}R_2 \to R_3$, we obtain

$$\rightsquigarrow \left[\begin{array}{ccc|ccc} \overline{1} & \overline{0} & \overline{3} & \overline{4} & \overline{0} & \overline{6} \\ \overline{0} & \overline{2} & \overline{0} & \overline{4} & \overline{0} & \overline{1} \\ \overline{0} & \overline{0} & \overline{2} & \overline{0} & \overline{1} & \overline{4} \end{array}\right].$$

From $4R_2 \to R_2$ and $4R_3 \to R_3$, we obtain

$$\rightsquigarrow \left[\begin{array}{ccc|ccc} \overline{1} & \overline{0} & \overline{3} & \overline{4} & \overline{0} & \overline{6} \\ \overline{0} & \overline{1} & \overline{0} & \overline{2} & \overline{0} & \overline{4} \\ \overline{0} & \overline{0} & \overline{1} & \overline{0} & \overline{4} & \overline{2} \end{array}\right].$$

Finally, from $R_1 - \overline{3}R_3 \to R_3$ we obtain

$$\rightsquigarrow \left[\begin{array}{ccc|ccc} \overline{1} & \overline{0} & \overline{0} & \overline{4} & \overline{2} & \overline{0} \\ \overline{0} & \overline{1} & \overline{0} & \overline{2} & \overline{0} & \overline{4} \\ \overline{0} & \overline{0} & \overline{1} & \overline{0} & \overline{4} & \overline{2} \end{array}\right].$$

We deduce that

$$\begin{bmatrix} \overline{1} & \overline{2} & \overline{3} \\ \overline{2} & \overline{3} & \overline{1} \\ \overline{3} & \overline{1} & \overline{2} \end{bmatrix}^{-1} = \begin{bmatrix} \overline{4} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{4} \\ \overline{0} & \overline{4} & \overline{2} \end{bmatrix}.$$

2.8.2 To find the inverse of B we consider the augmented matrix

$$\left[\begin{array}{ccc|ccc} 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & -1 & a & 0 & 1 & 0 \\ -1 & a & b & 0 & 0 & 1 \end{array}\right].$$

Performing $R_1 \leftrightarrow R_3$, $-R_3 \to R_3$, in succession,

$$\left[\begin{array}{ccc|ccc} -1 & a & b & 0 & 0 & 1 \\ 0 & -1 & a & 0 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{array}\right].$$

Performing $R_1 + aR_2 \to R_1$ and $R_2 - aR_3 \to R_2$ in succession,

$$\left[\begin{array}{ccc|ccc} -1 & 0 & b+a^2 & 0 & a & 1 \\ 0 & -1 & 0 & a & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{array}\right].$$

Performing $R_1 - (b+a^2)R_3 \to R_3$, $-R_1 \to R_1$ and $-R_2 \to R_2$ in succession, we find

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -b-a^2 & -a & -1 \\ 0 & 1 & 0 & -a & -1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{array}\right],$$

whence

$$B^{-1} = \begin{bmatrix} -b-a^2 & -a & -1 \\ -a & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix}.$$

Now,

$$\begin{aligned}
BAB^{-1} &= \begin{bmatrix} 0 & 0 & -1 \\ 0 & -1 & a \\ -1 & a & b \end{bmatrix} \begin{bmatrix} a & b & c \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -b-a^2 & -a & -1 \\ -a & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & -1 & 0 \\ -1 & a & 0 \\ 0 & 0 & -c \end{bmatrix} \begin{bmatrix} -b-a^2 & -a & -1 \\ -a & -1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} a & 1 & 0 \\ b & 0 & 1 \\ c & 0 & 0 \end{bmatrix} \\
&= A^T,
\end{aligned}$$

which is what we wanted to prove.

2.8.3 First, form the augmented matrix:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & x & 0 & 0 & 1 \end{array}\right].$$

Perform $R_2 - R_1 \to R_2$ and $R_3 - R_1 \to R_3$:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 1 & x & -1 & 0 & 1 \end{array}\right].$$

Performing $R_3 - R_2 \to R_3$:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & x & 0 & -1 & 1 \end{array}\right].$$

Finally, performing $\frac{1}{x}R_3 \to R_3$:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{x} & \frac{1}{x} \end{array}\right].$$

2.8.4 Since $MM^{-1} = \mathbf{I}_3$, multiplying the first row of M times the third column of M^{-1}, and again, the third row of M times the third column of M^{-1}, we gather that

$$1 \cdot 0 + 0 \cdot a + 1 \cdot b = 0, \qquad 0 \cdot 0 + 1 \cdot a + 1 \cdot b = 1 \implies b = 0, a = 1.$$

2.8.5 It is easy to prove by induction that $A^n = \begin{bmatrix} 1 & 0 & 0 \\ n & 1 & 0 \\ \frac{n(n+1)}{2} & n & 1 \end{bmatrix}$. Row-reducing, $(A^n)^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ -n & 1 & 0 \\ \frac{(n-1)n}{2} & -n & 1 \end{bmatrix}$.

2.8.6 Take, for example, $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = -A^{-1}$.

2.8.7 Operating formally, and using elementary row operations, we find

$$B^{-1} = \begin{bmatrix} -\frac{a^2-1}{a^2-5+2a} & \frac{a^2+2a-2}{a^2-5+2a} & \frac{a-2}{a^2-5+2a} \\ -\frac{2}{a^2-5+2a} & \frac{a+4}{a^2-5+2a} & -\frac{1}{a^2-5+2a} \\ \frac{2a}{a^2-5+2a} & -\frac{2a+5}{a^2-5+2a} & \frac{a}{a^2-5+2a} \end{bmatrix}.$$

Thus B is invertible whenever $a \neq -1 \pm \sqrt{6}$.

2.8.8 Form the augmented matrix

$$\left[\begin{array}{ccc|ccc} a & 2a & 3a & 1 & 0 & 0 \\ 0 & b & 2b & 0 & 1 & 0 \\ 0 & 0 & c & 0 & 0 & 1 \end{array}\right].$$

Perform $\frac{1}{a}R_1 \to R_1$, $\frac{1}{b}R_2 \to R_2$, $\frac{1}{a}R_3 \to R_3$, in succession, obtaining

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1/a & 0 & 0 \\ 0 & 1 & 2 & 0 & 1/b & 0 \\ 0 & 0 & 1 & 0 & 0 & 1/c \end{array}\right].$$

Now perform $R_1 - 2R_2 \to R_1$ and $R_2 - 2R_3 \to R_2$ in succession, to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 0 & -1 & 1/a & -2/a & 0 \\ 0 & 1 & 0 & 0 & 1/b & -2/c \\ 0 & 0 & 1 & 0 & 0 & 1/c \end{array}\right].$$

Finally, perform $R_1 + R_3 \to R_1$ to obtain

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/a & -2/b & 1/c \\ 0 & 1 & 0 & 0 & 1/b & -2/c \\ 0 & 0 & 1 & 0 & 0 & 1/c \end{array}\right].$$

Whence

$$\begin{bmatrix} a & 2a & 3a \\ 0 & b & 2b \\ 0 & 0 & c \end{bmatrix}^{-1} = \begin{bmatrix} 1/a & -2/b & 1/c \\ 0 & 1/b & -2/c \\ 0 & 0 & 1/c \end{bmatrix}.$$

2.8.9 To compute the inverse matrix we proceed formally as follows. The augmented matrix is

$$\left[\begin{array}{ccc|ccc} b & a & 0 & 1 & 0 & 0 \\ c & 0 & a & 0 & 1 & 0 \\ 0 & c & b & 0 & 0 & 1 \end{array}\right].$$

Performing $bR_2 - cR_1 \to R_2$ we find

$$\left[\begin{array}{ccc|ccc} b & a & 0 & 1 & 0 & 0 \\ 0 & -ca & ab & -c & b & 0 \\ 0 & c & b & 0 & 0 & 1 \end{array}\right].$$

Performing $aR_3 + R_2 \to R_3$ we obtain

$$\left[\begin{array}{ccc|ccc} b & a & 0 & 1 & 0 & 0 \\ 0 & -ca & ab & -c & b & 0 \\ 0 & 0 & 2ab & -c & b & a \end{array}\right].$$

Performing $2R_2 - R_3 \to R_2$ we obtain

$$\left[\begin{array}{ccc|ccc} b & a & 0 & 1 & 0 & 0 \\ 0 & -2ca & 0 & -c & b & -a \\ 0 & 0 & 2ab & -c & b & a \end{array}\right].$$

Performing $2cR_1 + R_2 \to R_1$ we obtain

$$\left[\begin{array}{ccc|ccc} 2bc & 0 & 0 & c & b & -a \\ 0 & -2ca & 0 & -c & b & -a \\ 0 & 0 & 2ab & -c & b & a \end{array}\right].$$

From here we easily conclude that

$$\begin{bmatrix} b & a & 0 \\ c & 0 & a \\ 0 & c & b \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2b} & \frac{1}{2c} & -\frac{a}{2bc} \\ \frac{1}{2a} & -\frac{b}{2ac} & \frac{1}{2c} \\ -\frac{c}{2ba} & \frac{1}{2a} & \frac{1}{2b} \end{bmatrix}$$

as long as $abc \neq 0$.

2.8.10 Since AB is invertible, $\textbf{rank}(AB) = n$. Thus

$$n = \textbf{rank}(AB) \leq \textbf{rank}(A) \leq n \implies \textbf{rank}(A) = n,$$

$$n = \textbf{rank}(AB) \leq \textbf{rank}(B) \leq n \implies \textbf{rank}(B) = n,$$

whence A and B are invertible.

2.8.11 Form the expanded matrix

$$\left[\begin{array}{ccc|ccc} 1+a & 1 & 1 & 1 & 0 & 0 \\ 1 & 1+b & 1 & 0 & 1 & 0 \\ 1 & 1 & 1+c & 0 & 0 & 1 \end{array}\right].$$

Perform $bcR_1 \to R_1$, $abR_3 \to R_3$, $caR_2 \to R_2$. The matrix turns into

$$\left[\begin{array}{ccc|ccc} bc+abc & bc & bc & bc & 0 & 0 \\ ca & ca+abc & ca & 0 & ca & 0 \\ ab & ab & ab+abc & 0 & 0 & ab \end{array}\right].$$

Perform $R_1 + R_2 + R_3 \to R_1$ the matrix turns into

$$\left[\begin{array}{ccc|ccc} ab+bc+ca+abc & ab+bc+ca+abc & ab+bc+ca+abc & bc & ca & ab \\ ca & ca+abc & ca & 0 & ca & 0 \\ ab & ab & ab+abc & 0 & 0 & ab \end{array}\right].$$

Perform $\frac{1}{ab+bc+ca+abc}R_1 \to R_1$. The matrix turns into

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & \frac{bc}{ab+bc+ca+abc} & \frac{ca}{ab+bc+ca+abc} & \frac{ab}{ab+bc+ca+abc} \\ ca & ca+abc & ca & 0 & ca & 0 \\ ab & ab & ab+abc & 0 & 0 & ab \end{array}\right].$$

Perform $R_2 - caR_1 \to R_2$ and $R_3 - abR_3 \to R_3$. We get

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & \frac{bc}{ab+bc+ca+abc} & \frac{ca}{ab+bc+ca+abc} & \frac{ab}{ab+bc+ca+abc} \\ 0 & abc & 0 & -\frac{abc^2}{ab+bc+ca+abc} & ca-\frac{c^2a^2}{ab+bc+ca+abc} & -\frac{a^2bc}{ab+bc+ca+abc} \\ 0 & 0 & abc & -\frac{ab^2c}{ab+bc+ca+abc} & -\frac{a^2bc}{ab+bc+ca+abc} & ab-\frac{a^2b^2}{ab+bc+ca+abc} \end{array}\right].$$

Perform $\frac{1}{ABC}R_2 \to R_2$ and $\frac{1}{ABC}R_3 \to R_3$. We obtain

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & \frac{bc}{ab+bc+ca+abc} & \frac{ca}{ab+bc+ca+abc} & \frac{ab}{ab+bc+ca+abc} \\ 0 & 1 & 0 & -\frac{c}{ab+bc+ca+abc} & \frac{1}{b}-\frac{ca}{b(ab+bc+ca+abc)} & -\frac{a}{ab+bc+ca+abc} \\ 0 & 0 & 1 & -\frac{b}{ab+bc+ca+abc} & -\frac{a}{ab+bc+ca+abc} & \frac{1}{c}-\frac{ab}{c(ab+bc+ca+abc)} \end{array}\right].$$

Finally we perform $R_1 - R_2 - R_3 \to R_1$, getting

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{a+b+bc}{ab+bc+ca+abc} & -\frac{c}{ab+bc+ca+abc} & -\frac{b}{ab+bc+ca+abc} \\ 0 & 1 & 0 & -\frac{c}{ab+bc+ca+abc} & \frac{1}{b}-\frac{ca}{b(ab+bc+ca+abc)} & -\frac{a}{ab+bc+ca+abc} \\ 0 & 0 & 1 & -\frac{b}{ab+bc+ca+abc} & -\frac{a}{ab+bc+ca+abc} & \frac{1}{c}-\frac{ab}{c(ab+bc+ca+abc)} \end{array}\right].$$

We conclude that the inverse is

$$\begin{bmatrix} \frac{b+c+bc}{ab+bc+ca+abc} & -\frac{c}{ab+bc+ca+abc} & -\frac{b}{ab+bc+ca+abc} \\ -\frac{c}{ab+bc+ca+abc} & \frac{c+a+ca}{ab+bc+ca+abc} & -\frac{a}{ab+bc+ca+abc} \\ -\frac{b}{ab+bc+ca+abc} & -\frac{a}{ab+bc+ca+abc} & \frac{a+b+ab}{ab+bc+ca+abc} \end{bmatrix}$$

2.8.16 Since $\textbf{rank}\left(A^2\right) < 5$, A^2 is not invertible. But then A is not invertible and hence $\textbf{rank}\left(A\right) < 5$.

2.8.17 Each entry can be chosen in p ways, which means that there are p^2 ways of choosing the two entries of an arbitrary row. The first row cannot be the zero row, hence there are $p^2 - 1$ ways of choosing it. The second row cannot be one of the p multiples of the first row, hence there are $p^2 - p$ ways of choosing it. In total, this gives $(p^2 - 1)(p^2 - p)$ invertible matrices in $\mathbb{Z}_p$.

2.8.18 Assume that both A and B are $m \times n$ matrices. Let $C = [A \quad B]$ be the $m \times (2n)$ obtained by juxtaposing A to B. $\textbf{rank}\left(C\right)$ is the number of linearly independent columns of C, which is composed of the columns of A and B. By column-reducing the first n columns, we find $\textbf{rank}\left(A\right)$ linearly independent columns. By column-reducing columns $n + 1$ to $2n$, we find $\textbf{rank}\left(B\right)$ linearly independent columns. These $\textbf{rank}\left(A\right) + \textbf{rank}\left(B\right)$ columns are distinct, and are a subset of the columns of C. Since C has at most $\textbf{rank}\left(C\right)$ linearly independent columns, it follows that $\textbf{rank}\left(C\right) \leq \textbf{rank}\left(A\right) + \textbf{rank}\left(B\right)$. Furthermore, by adding the $n + k$-column $(1 \leq k \leq n)$ of C to the k-th column, we see that C is column-equivalent to $[A + B \ B]$. But clearly

$$\textbf{rank}\left(A + B\right) \leq \textbf{rank}\left([A + B \ B]\right) = \textbf{rank}\left(C\right),$$

since $[A + B \ B]$ is obtained by adding columns to $A + B$. We deduce

$$\textbf{rank}\left(A + B\right) \leq \textbf{rank}\left([A + B \ B]\right) = \textbf{rank}\left(C\right) \leq \textbf{rank}\left(A\right) + \textbf{rank}\left(B\right),$$

as was to be shewn.

2.8.19 Since the first two columns of AB are not proportional, and since the last column is the sum of the first two, $\textbf{rank}\left(AB\right) = 2$. Now,

$$(AB)^2 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix}^2 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{bmatrix} = AB.$$

Since BA is a 2×2 matrix, $\textbf{rank}\left(BA\right) \leq 2$. Also,

$$2 = \textbf{rank}\left(AB\right) = \textbf{rank}\left((AB)^2\right) = \textbf{rank}\left(A(BA)B\right) \leq \textbf{rank}\left(BA\right),$$

whence $\textbf{rank}\left(BA\right) = 2$, which means BA is invertible. Finally,

$$(AB)^2 - AB = \mathbf{0}_3 \implies A(BA - \mathbf{I}_2)B = \mathbf{0}_3 \implies BA(BA - \mathbf{I}_2)BA = B\mathbf{0}_3A \implies BA - \mathbf{I}_2 = \mathbf{0}_2,$$

since BA is invertible and we may cancel it.

3.1.1 The free variables are z and w. We have

$$\overline{2}y + w = \overline{2} \implies \overline{2}y = \overline{2} - w \implies y = \overline{1} + w,$$

and

$$x + y + z + w = \overline{0} \implies x = -y - z - w = \overline{2}y + \overline{2}z + \overline{2}w.$$

Hence

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \\ \overline{0} \end{bmatrix} + z \begin{bmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \\ \overline{0} \end{bmatrix} + w \begin{bmatrix} \overline{0} \\ \overline{0} \\ \overline{0} \\ \overline{1} \end{bmatrix}.$$

This gives the 9 solutions.

3.1.2 We have

$$\begin{bmatrix} \overline{1} & \overline{2} & \overline{3} \\ \overline{2} & \overline{3} & \overline{1} \\ \overline{3} & \overline{1} & \overline{2} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{5} \\ \overline{6} \\ \overline{0} \end{bmatrix},$$

Hence

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \overline{1} & \overline{2} & \overline{3} \\ \overline{2} & \overline{3} & \overline{1} \\ \overline{3} & \overline{1} & \overline{2} \end{bmatrix}^{-1} \begin{bmatrix} \overline{5} \\ \overline{6} \\ \overline{0} \end{bmatrix} = \begin{bmatrix} \overline{4} & \overline{2} & \overline{0} \\ \overline{2} & \overline{0} & \overline{4} \\ \overline{0} & \overline{4} & \overline{2} \end{bmatrix} \begin{bmatrix} \overline{5} \\ \overline{6} \\ \overline{0} \end{bmatrix} = \begin{bmatrix} \overline{4} \\ \overline{3} \\ \overline{3} \end{bmatrix}.$$

3.1.3 The augmented matrix of the system is

$$\left[\begin{array}{ccc|c} \overline{1} & -\overline{2} & \overline{1} & \overline{5} \\ \overline{2} & \overline{2} & \overline{0} & \overline{7} \\ \overline{5} & -\overline{3} & \overline{4} & \overline{1} \end{array}\right] \underset{R_2 - \overline{2}R_1 \to R_2}{\overset{R_3 - \overline{5}R_1 \to R_3}{\rightsquigarrow}} \left[\begin{array}{ccc|c} \overline{1} & -\overline{2} & \overline{1} & \overline{5} \\ \overline{0} & \overline{6} & -\overline{2} & -\overline{3} \\ \overline{0} & \overline{7} & -\overline{1} & \overline{2} \end{array}\right]$$

$$\overset{\overline{6}R_3 - \overline{7}R_2 \to R_3}{\rightsquigarrow} \left[\begin{array}{ccc|c} \overline{1} & -\overline{2} & \overline{1} & \overline{5} \\ \overline{0} & \overline{6} & -\overline{2} & -\overline{3} \\ \overline{0} & \overline{0} & \overline{8} & \overline{7} \end{array}\right]$$

Backward substitution yields

$$\overline{8}z = \overline{7} \implies \overline{5} \cdot \overline{8}z = \overline{5} \cdot \overline{7} \implies z = \overline{35} = \overline{9},$$
$$\overline{6}y = \overline{2}z - \overline{3} = \overline{2} \cdot \overline{9} - \overline{3} = \overline{15} = \overline{2} \implies \overline{11} \cdot \overline{6}y = \overline{11} \cdot \overline{2} \implies y = \overline{22} = \overline{9},$$
$$x = \overline{2}y - \overline{1}z + \overline{5} = \overline{2} \cdot \overline{9} - \overline{1} \cdot \overline{9} + \overline{5} = \overline{14} = \overline{1}.$$

Conclusion :

$$\boxed{x = \overline{1}, \qquad y = \overline{9}, \qquad z = \overline{9}.}$$

Check:

$$\overline{1} - \overline{2} \cdot \overline{9} + \overline{9} = -\overline{8} \overset{\checkmark}{=} \overline{5},$$
$$\overline{2} \cdot \overline{1} + \overline{2} \cdot \overline{9} = \overline{20} \overset{\checkmark}{=} \overline{7},$$
$$\overline{5} \cdot \overline{1} - \overline{3} \cdot \overline{9} + \overline{4} \cdot \overline{9} = \overline{14} \overset{\checkmark}{=} \overline{1}.$$

3.1.4 We need to solve the system

$$a - b + c - d = p(-1) = -10,$$
$$a = p(0) = -1,$$
$$a + b + c + d = p(1) = 2,$$
$$a + 2b + 4c + 8d = p(2) = 23.$$

Using row reduction or otherwise, we find $a = -1$, $b = 2$, $c = -3$, $d = 4$, and so the polynomial is

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

3.1.5 Using the encoding chart

0	1	2	3	4	5	6	7	8	9	10	11	12
A	B	C	D	E	F	G	H	I	J	K	L	M
13	14	15	16	17	18	19	20	21	22	23	24	25
N	O	P	Q	R	S	T	U	V	W	X	Y	Z

we find

$$P_2 = \begin{bmatrix} \mathbf{M} \\ \mathbf{U} \\ \mathbf{N} \end{bmatrix} = \begin{bmatrix} 12 \\ 20 \\ 13 \end{bmatrix}, \quad P_3 = \begin{bmatrix} \mathbf{I} \\ \mathbf{S} \\ \mathbf{T} \end{bmatrix} = \begin{bmatrix} 8 \\ 18 \\ 19 \end{bmatrix}, \quad P_4 = \begin{bmatrix} \mathbf{S} \\ \mathbf{E} \\ \mathbf{A} \end{bmatrix} = \begin{bmatrix} 18 \\ 4 \\ 0 \end{bmatrix}, \quad P_5 = \begin{bmatrix} \mathbf{T} \\ \mathbf{O} \\ \mathbf{F} \end{bmatrix} = \begin{bmatrix} 19 \\ 14 \\ 5 \end{bmatrix}, \quad P_6 = \begin{bmatrix} \mathbf{F} \\ \mathbf{A} \\ \mathbf{L} \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \\ 11 \end{bmatrix}.$$

Thus

$$AP_2 = \begin{bmatrix} 20 \\ 10 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{U} \\ \mathbf{K} \\ \mathbf{A} \end{bmatrix}, \quad AP_3 = \begin{bmatrix} 18 \\ 24 \\ 12 \end{bmatrix} = \begin{bmatrix} \mathbf{S} \\ \mathbf{Y} \\ \mathbf{M} \end{bmatrix}, \quad AP_4 = \begin{bmatrix} 4 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{E} \\ \mathbf{C} \\ \mathbf{A} \end{bmatrix}, \quad AP_5 = \begin{bmatrix} 14 \\ 5 \\ 10 \end{bmatrix} = \begin{bmatrix} \mathbf{O} \\ \mathbf{F} \\ \mathbf{K} \end{bmatrix}, \quad AP_6 = \begin{bmatrix} 0 \\ 15 \\ 22 \end{bmatrix} = \begin{bmatrix} \mathbf{A} \\ \mathbf{P} \\ \mathbf{W} \end{bmatrix}.$$

Finally, the message is encoded into

OGY UKA SYM ECA OFK APW.

3.1.6 Observe that since 103 is prime, $\mathbb{Z}_{103}$ is a field. Adding the first hundred equations,

$$100x_0 + x_1 + x_2 + \cdots + x_{100} = 4950 \implies 99x_0 = 4950 - 4949 = 1 \implies x_0 = 77 \quad \mathbf{mod}\ 103.$$

Now, for $1 \leq k \leq 100$,

$$x_k = k - 1 - x_0 = k - 78 = k + 25.$$

This gives

$$x_1 = 26, x_2 = 27, \ldots, x_{77} = 102, x_{78} = 0, x_{79} = 1, x_{80} = 2, \ldots, x_{100} = 22.$$

3.3.1 Observe that the third row is the sum of the first two rows and the fourth row is twice the third. So we have

$$\left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & -1 \\ 2 & 1 & 2 & 1 & 2 & 0 \\ 4 & 2 & 4 & 2 & 4 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right] \underset{R_4 - 2R_1 - 2R_2 \to R_4}{\overset{R_3 - R_1 - R_2 \to R_3}{\rightsquigarrow}} \left[\begin{array}{ccccc|c} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

$$\underset{R_1 - R_5 \to R_1}{\overset{R_2 - R_5 \to R_2}{\rightsquigarrow}} \left[\begin{array}{ccccc|c} 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{array}\right]$$

Rearranging the rows we obtain

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Hence d and f are free variables. We obtain

$$c = -1,$$

$$b = 1 - c - d = 2 - d,$$

$$a = -f.$$

The solution is

$$\begin{bmatrix} a \\ b \\ c \\ d \\ f \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ -1 \\ 0 \\ 0 \end{bmatrix} + d \begin{bmatrix} 0 \\ -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + f \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

3.3.2 The unique solution is $\begin{bmatrix} 1 \\ 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$.

3.3.3 The augmented matrix of the system is

$$\left[\begin{array}{ccc|c} 2m & 1 & 1 & 2 \\ 1 & 2m & 1 & 4m \\ 1 & 1 & 2m & 2m^2 \end{array}\right].$$

Performing $R_1 \leftrightarrow R_2$.

$$\left[\begin{array}{ccc|c} 1 & 2m & 1 & 4m \\ 2m & 1 & 1 & 2 \\ 1 & 1 & 2m & 2m^2 \end{array}\right].$$

Performing $R_2 \leftrightarrow R_3$.

$$\left[\begin{array}{ccc|c} 1 & 2m & 1 & 4m \\ 1 & 1 & 2m & 2m^2 \\ 2m & 1 & 1 & 2 \end{array}\right].$$

Performing $R_2 - R_1 \to R_1$ and $R_3 - 2mR_1 \to R_3$ we obtain

$$\left[\begin{array}{ccc|c} 1 & 2m & 1 & 4m \\ 0 & 1-2m & 2m-1 & 2m^2-4m \\ 0 & 1-4m^2 & 1-2m & 2-8m^2 \end{array}\right].$$

If $m = \frac{1}{2}$ the matrix becomes

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & -\frac{3}{2} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

and hence it does not have a solution. If $m \neq \frac{1}{2}$, by performing $\frac{1}{1-2m}R_2 \to R_2$ and $\frac{1}{1-2m}R_3 \to R_3$, the matrix becomes

$$\left[\begin{array}{ccc|c} 1 & 2m & 1 & 4m \\ 0 & 1 & -1 & \frac{2m(m-2)}{1-2m} \\ 0 & 1+2m & 1 & 2(1+2m) \end{array}\right].$$

Performing $R_3 - (1+2m)R_2 \to R_3$ we obtain

$$\left[\begin{array}{ccc|c} 1 & 2m & 1 & 4m \\ 0 & 1 & -1 & \frac{2m(m-2)}{1-2m} \\ 0 & 0 & 2+2m & \frac{2(1+2m)(1-m^2)}{1-2m} \end{array}\right].$$

If $m = -1$ then the matrix reduces to

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & -4 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

The solution in this case is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} z \\ 2+z \\ z \end{bmatrix}.$$

If $m \neq -1, m \neq -\frac{1}{2}$ we have the solutions

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{m-1}{1-2m} \\ \frac{1-3m}{1-2m} \\ \frac{(1+2m)(1-m)}{1-2m} \end{bmatrix}.$$

3.3.4 By performing the elementary row operations, we obtain the following triangular form:

$$ax + y - 2z = 1,$$
$$(a-1)^2 y + (1-a)(a-2)z = 1-a,$$
$$(a-2)z = 0.$$

If $a = 2$, there is an infinity of solutions:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1+t \\ -1 \\ t \end{bmatrix} \quad t \in \mathbb{R}.$$

Assume $a \neq 2$. Then $z = 0$ and the system becomes

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

We see that if $a = 1$, the system becomes

$$\begin{aligned} x + y &= 1, \\ 2x + 2y &= 1, \end{aligned}$$

and so there is no solution. If $(a-1)(a-2) \neq 0$, we obtain the unique solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{a-1} \\ -\frac{1}{a-1} \\ 0 \end{bmatrix}.$$

3.3.5 The system is solvable if $m \neq 0, m \neq \pm 2$. If $m \neq 2$ there is the solution $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \frac{1}{m-2} \\ \frac{m+3}{m-2} \\ \frac{m+2}{m-2} \end{bmatrix}$.

3.3.6 There is the unique solution $\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} a+d+b-c \\ -c-d-b+a \\ d+c-b+a \\ c-d+b+a \end{bmatrix}$.

3.3.7 The system can be written as

$$\begin{bmatrix} b & a & 0 \\ c & 0 & a \\ 0 & c & b \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} c \\ b \\ a \end{bmatrix}.$$

The system will have the unique solution

$$\begin{aligned} \begin{bmatrix} x \\ y \\ z \end{bmatrix} &= \begin{bmatrix} b & a & 0 \\ c & 0 & a \\ 0 & c & b \end{bmatrix}^{-1} \begin{bmatrix} c \\ b \\ a \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{2b} & \frac{1}{2c} & -\frac{a}{2bc} \\ \frac{1}{2a} & -\frac{b}{2ac} & \frac{1}{2c} \\ -\frac{c}{2ba} & \frac{1}{2a} & \frac{1}{2b} \end{bmatrix} \begin{bmatrix} c \\ b \\ a \end{bmatrix}, \\ &= \begin{bmatrix} \dfrac{b^2+c^2-a^2}{2bc} \\ \dfrac{a^2+c^2-b^2}{2ac} \\ \dfrac{a^2+b^2-c^2}{2ab} \end{bmatrix} \end{aligned}$$

as long as the inverse matrix exists, which is as long as $abc \neq 0$

3.3.8 We first form the augmented matrix,

$$\left[\begin{array}{ccc|c} 1-a & 2a+1 & 2a+2 & a \\ a & a & 0 & 2a+2 \\ 2 & a+1 & a-1 & a^2-2a+9 \end{array}\right] \overset{R_1+R_2\to R_1}{\rightsquigarrow} \left[\begin{array}{ccc|c} 1 & 3a+1 & 2a+2 & 3a+2 \\ a & a & 0 & 2a+2 \\ 2 & a+1 & a-1 & a^2-2a+9 \end{array}\right]$$

$$\underset{R_3-2R_1\to R_3}{\overset{R_2-aR_1\to R_2}{\rightsquigarrow}} \left[\begin{array}{ccc|c} 1 & 3a+1 & 2a+2 & 3a+2 \\ 0 & -3a^2 & -2a^2-2a & -3a^2+2 \\ 0 & -5a-1 & -3a-5 & a^2-8a+5 \end{array}\right].$$

After $(-5a-1)R_2 + 3a^2R_3 \to R_2$, this las matrix becomes

$$\left[\begin{array}{ccc|c} 1 & 3a+1 & 2a+2 & 3a+2 \\ 0 & 0 & a^3-3a^2+2a & 3a^4-9a^3+18a^2-10a-2 \\ 0 & -5a-1 & -3a-5 & a^2-8a+5 \end{array}\right].$$

Exchanging the last two rows and factoring,

$$\left[\begin{array}{ccc|c} 1 & 3a+1 & 2a+2 & 3a+2 \\ 0 & -5a-1 & -3a-5 & a^2-8a+5 \\ 0 & 0 & a(a-1)(a-2) & (a-1)(3a^3-6a^2+12a+2) \end{array}\right].$$

Thus we must examine $a \in \{1, 2, 3\}$ and $a \not\in \{0, 1, 2\}$.

Clearly, if $a(a-1)(a-2) \neq 0$, then there is the unique solution

$$\left\{z = \frac{2+12a-6a^2+3a^3}{a\,(a-2)}, \qquad y = -\frac{2a^3-3\,a^2+6a+10}{a\,(a-2)}, \qquad x = \frac{2\,a^3-a^2+4a+6}{a\,(a-2)}\right\}.$$

If $a = 0$, the system becomes

$$x + y + 2z = 0, \quad 0 = 2, \quad 2x + y - z = 0,$$

which is inconsistent (no solutions).

If $a = 1$, the system becomes

$$3y + 4z = 1, \qquad x + y = 1, \qquad 2x + 2y = 8,$$

which has infinitely many solutions,

$$\left\{y = \frac{1}{3} - \frac{4}{3}z, \qquad x = \frac{2}{3} + \frac{4}{3}z, \qquad z = z\right\}.$$

If $a = 2$, the system becomes

$$-x + 5y + 6z = 2, \quad 2x + 2y = 6, \qquad 2x + 3y + z = 9,$$

which is also inconsistent, as can be seen by observing that

$$(-x + 5y + 6z) - 6(2x + 3y + z) = 2 - 18 \implies -13x - 13y = -18,$$

which contradicts the equation $2x + 2y = 6$.

3.3.9

$$\begin{aligned} x &= 2^{-2}3^{6} \\ y &= 2^{-3}3^{12} \\ z &= 2^{2}3^{-7}. \end{aligned}$$

3.3.10 Denote the addition operations applied to the rows by a_1, a_2, a_3, a_4 and the subtraction operations to the columns by b_1, b_2, b_3, b_4. Comparing A and A^T we obtain 7 equations in 8 unknowns. By inspecting the diagonal entries, and the entries of the first row of A and A^T, we deduce the following equations

$$\begin{aligned} a_1 &= b_1, \\ a_2 &= b_2, \\ a_3 &= b_3, \\ a_4 &= b_4, \\ a_1 - b_2 &= 3, \\ a_1 - b_3 &= 6, \\ a_1 - b_4 &= 9. \end{aligned}$$

This is a system of 7 equations in 8 unknowns. We may let $a_4 = 0$ and thus obtain $a_1 = b_1 = 9$, $a_2 = b_2 = 6$, $a_3 = b_3 = 3$, $a_4 = b_4 = 0$.

3.3.11 The augmented matrix of this system is

$$\left[\begin{array}{ccccc|c} -y & 1 & 0 & 0 & 1 & 0 \\ 1 & -y & 1 & 0 & 0 & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 1 & 0 & 0 & 1 & -y & 0 \end{array}\right].$$

Permute the rows to obtain

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 1 & -y & 1 & 0 & 0 & 0 \\ -y & 1 & 0 & 0 & 1 & 0 \end{array}\right].$$

Performing $R_5 + yR_1 \to R_5$ and $R_4 - R_1 \to R_4$ we get

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 0 & -y & 1 & -1 & y & 0 \\ 0 & 1 & 0 & y & 1-y^2 & 0 \end{array}\right].$$

Performing $R_5 - R_2 \to R_5$ and $R_4 + yR_2 \to R_4$ we get

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 0 & 0 & 1-y^2 & y-1 & y & 0 \\ 0 & 0 & y & y-1 & 1-y^2 & 0 \end{array}\right].$$

Performing $R_5 - yR_3 \to R_5$ and $R_4 + (y^2-1)R_3 \to R_4$ we get

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 0 & 0 & 0 & -y^3+2y-1 & y^2+y-1 & 0 \\ 0 & 0 & 0 & y^2+y-1 & 1-y-y^2 & 0 \end{array}\right].$$

Performing $R_5 + R_4 \to R_5$ we get

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 0 & 0 & 0 & -y^3+2y-1 & y^2+y-1 & 0 \\ 0 & 0 & 0 & -y^3+y^2+3y-2 & 0 & 0 \end{array}\right].$$

Upon factoring, the matrix is equivalent to

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 0 & 0 & 0 & -(y-1)(y^2+y-1) & y^2+y-1 & 0 \\ 0 & 0 & 0 & -(y-2)(y^2+y-1) & 0 & 0 \end{array}\right].$$

Thus $(y-2)(y^2+y-1)x_4 = 0$. If $y = 2$ then the system reduces to

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -2 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & -5 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

In this case x_5 is free and by backwards substitution we obtain

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} t \\ t \\ t \\ t \\ t \end{bmatrix}, \quad t \in \mathbb{R}.$$

If $y^2 + y - 1 = 0$ then the system reduces to

$$\left[\begin{array}{ccccc|c} 1 & 0 & 0 & 1 & -y & 0 \\ 0 & 1 & -y & 1 & 0 & 0 \\ 0 & 0 & 1 & -y & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

In this case x_4, x_5 are free, and

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} yt - s \\ y^2s - yt - s \\ ys - t \\ s \\ t \end{bmatrix}, \quad (s, t) \in \mathbb{R}^2.$$

Since $y^2s - s = (y^2 + y - 1)s - ys$, this last solution can be also written as

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} yt - s \\ -ys - yt \\ ys - t \\ s \\ t \end{bmatrix}, \quad (s, t) \in \mathbb{R}^2.$$

Finally, if $(y - 2)(y^2 + y - 1) \neq 0$, then $x_4 = 0$, and we obtain

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

4.1.1 No, since $1_{\mathbb{F}}\overrightarrow{v} = \overrightarrow{v}$ is not fulfilled. For example

$$1 \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot 1 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

4.1.2 We expand $(1_{\mathbb{F}} + 1_{\mathbb{F}})(\overrightarrow{a} + \overrightarrow{b})$ in two ways, first using 4.7 first and then 4.8, obtaining

$$(1_{\mathbb{F}} + 1_{\mathbb{F}})(\overrightarrow{a} + \overrightarrow{b}) = (1_{\mathbb{F}} + 1_{\mathbb{F}})\overrightarrow{a} + (1_{\mathbb{F}} + 1_{\mathbb{F}})\overrightarrow{b} = \overrightarrow{a} + \overrightarrow{a} + \overrightarrow{b} + \overrightarrow{b},$$

and then using 4.8 first and then 4.7, obtaining

$$(1_{\mathbb{F}} + 1_{\mathbb{F}})(\overrightarrow{a} + \overrightarrow{b}) = 1_{\mathbb{F}}(\overrightarrow{a} + \overrightarrow{b}) + 1_{\mathbb{F}}(\overrightarrow{a} + \overrightarrow{b}) = \overrightarrow{a} + \overrightarrow{b} + \overrightarrow{a} + \overrightarrow{b}.$$

We thus have the equality

$$\overrightarrow{a} + \overrightarrow{a} + \overrightarrow{b} + \overrightarrow{b} = \overrightarrow{a} + \overrightarrow{b} + \overrightarrow{a} + \overrightarrow{b}.$$

Cancelling $\overrightarrow{a}$ from the left and $\overrightarrow{b}$ from the right, we obtain

$$\overrightarrow{a} + \overrightarrow{b} = \overrightarrow{b} + \overrightarrow{a},$$

which is what we wanted to shew.

4.1.3 We must prove that each of the axioms of a vector space are satisfied. Clearly if $(x, y, \alpha) \in \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}$ then $x \oplus y = xy > 0$ and $\alpha \otimes x = x^\alpha > 0$, so V is closed under vector addition and scalar multiplication. Commutativity and associativity of vector addition are obvious.

Let A be additive identity. Then we need

$$x \oplus A = x \implies xA = x \implies A = 1.$$

Thus the additive identity is 1. Suppose I is the additive inverse of x. Then

$$x \oplus I = 1 \implies xI = 1 \implies I = \frac{1}{x}.$$

Hence the additive inverse of x is $\frac{1}{x}$.

Now

$$\alpha \otimes (x \oplus y) = (xy)^\alpha = x^\alpha y^\alpha = x^\alpha \oplus y^\alpha = (\alpha \otimes x) \oplus (\alpha \otimes y),$$

and

$$(\alpha + \beta) \otimes x = x^{\alpha+\beta} = x^\alpha x^\beta = (x^\alpha) \oplus (x^\beta) = (\alpha \otimes x) \oplus (\beta \otimes x),$$

whence the distributive laws hold.

Finally,

$$1 \otimes x = x^1 = x,$$

and

$$\alpha \otimes (\beta \otimes x) = (\beta \otimes x)^\alpha = (x^\beta)^\alpha = x^{\alpha\beta} = (\alpha\beta) \otimes x,$$

and the last two axioms also hold.

4.1.4 $\mathbb{C}$ is a vector space over $\mathbb{R}$, the proof is trivial. But $\mathbb{R}$ is not a vector space over $\mathbb{C}$, since, for example taking i as a scalar (from $\mathbb{C}$) and 1 as a vector (from $\mathbb{R}$) the scalar multiple $i \cdot 1 = i \notin \mathbb{R}$ and so there is no closure under scalar multiplication.

4.1.5 One example is

$$(\mathbb{Z}_2)^3 = \left\{ \begin{bmatrix} \overline{0} \\ \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{0} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{1} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{1} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{0} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{1} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{1} \\ \overline{1} \end{bmatrix} \right\}.$$

Addition is the natural element-wise addition and scalar multiplication is ordinary element-wise scalar multiplication.

4.1.6 One example is

$$(\mathbb{Z}_3)^2 = \left\{ \begin{bmatrix} \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{2} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{2} \end{bmatrix}, \begin{bmatrix} \overline{2} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{2} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{2} \\ \overline{2} \end{bmatrix} \right\}.$$

Addition is the natural element-wise addition and scalar multiplication is ordinary element-wise scalar multiplication.

4.2.1 Take $\alpha \in \mathbb{R}$ and

$$\overrightarrow{x} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in X, \quad a - b - 3d = 0, \quad \overrightarrow{y} = \begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} \in X, \quad a' - b' - 3d' = 0.$$

Then

$$\overrightarrow{x} + \alpha\overrightarrow{y} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} + \alpha \begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} = \begin{bmatrix} a + \alpha a' \\ b + \alpha b' \\ c + \alpha c' \\ d + \alpha d' \end{bmatrix}.$$

Observe that

$$(a + \alpha a') - (b + \alpha b') - 3(d + \alpha d') = (a - b - 3d) + \alpha(a' - b' - 3d') = 0 + \alpha 0 = 0,$$

meaning that $\overrightarrow{x} + \alpha\overrightarrow{y} \in X$, and so X is a vector subspace of $\mathbb{R}^4$.

4.2.2 Take

$$\overrightarrow{u} = \begin{bmatrix} a_1 \\ 2a_1 - 3b_1 \\ 5b_1 \\ a_1 + 2b_1 \\ a_1 \end{bmatrix}, \overrightarrow{v} = \begin{bmatrix} a_2 \\ 2a_2 - 3b_2 \\ 5b_2 \\ a_2 + 2b_2 \\ a_2 \end{bmatrix}, \quad \alpha \in \mathbb{R}.$$

Put $s = a_1 + \alpha a_2$, $t = b_1 + \alpha b_2$. Then

$$\overrightarrow{u} + \alpha\overrightarrow{v} = \begin{bmatrix} a_1 + \alpha a_2 \\ 2(a_1 + \alpha a_2) - 3(b_1 + \alpha b_2) \\ 5(b_1 + \alpha b_2) \\ (a_1 + \alpha a_2) + 2(b_1 + \alpha b_2) \\ a_1 + \alpha a_2 \end{bmatrix} = \begin{bmatrix} s \\ 2s - 3t \\ 5t \\ s + 2t \\ s \end{bmatrix} \in X,$$

since this last matrix has the basic shape of matrices in X. This shews that X is a vector subspace of $\mathbb{R}^5$.

4.2.7 We shew that some of the properties in the definition of vector subspace fail to hold in these sets.

❶ Take $\vec{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\alpha = 2$. Then $\vec{x} \in V$ but $2\vec{x} = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \not\in V$ as $0^2 + 2^2 = 4 \neq 1$. So V is not closed under scalar multiplication.

❷ Take $\vec{x} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$. Then $\vec{x} \in W, \vec{y} \in W$ but $\vec{x} + \vec{y} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \not\in W$ as $1 \cdot 1 = 1 \neq 0$. Hence W is not closed under vector addition.

❸ Take $\vec{x} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$. Then $\vec{x} \in Z$ but $-\vec{x} = -\begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \not\in Z$ as $1 + (-1)^2 = 2 \neq 0$. So Z is not closed under scalar multiplication.

4.2.8 Assume $U_1 \not\subseteq U_2$ and $U_2 \not\subseteq U_1$. Take $\vec{v} \in U_2 \setminus U_1$ (which is possible because $U_2 \not\subseteq U_1$) and $\vec{u} \in U_1 \setminus U_2$ (which is possible because $U_1 \not\subseteq U_2$). If $\vec{u} + \vec{v} \in U_1$, then—as $-\vec{u}$ is also in U_1—the sum of two vectors in U_1 must also be in U_1 giving

$$\vec{u} + \vec{v} - \vec{u} = \vec{v} \in U_1,$$

a contradiction. Similarly if $\vec{u} + \vec{v} \in U_2$, then—as $-\vec{v}$ also in U_2—the sum of two vectors in U_2 must also be in U_1 giving

$$\vec{u} + \vec{v} - \vec{u} = \vec{u} \in U_2,$$

another contradiction. Hence either $U_1 \subseteq U_2$ or $U_2 \subseteq U_1$ (or possibly both).

4.2.9 Assume contrariwise that $V = U_1 \bigcup U_2 \bigcup \cdots \bigcup U_k$ is the shortest such list. Since the U_j are proper subspaces, $k > 1$. Choose $\vec{x} \in U_1, \vec{x} \not\in U_2 \bigcup \cdots \bigcup U_k$ and choose $\vec{y} \not\in U_1$. Put $L = \{\vec{y} + \alpha\vec{x} | \alpha \in \mathbb{F}\}$. Claim: $L \bigcap U_1 = \varnothing$. For if $\vec{u} \in L \bigcap U_1$ then $\exists a_0 \in \mathbb{F}$ with $\vec{u} = \vec{y} + a_0\vec{x}$ and so $\vec{y} = \vec{u} - a_0\vec{x} \in U_1$, a contradiction. So L and U_1 are disjoint.

We now shew that L has at most one vector in common with $U_j, 2 \leq j \leq k$. For, if there were two elements of $\mathbb{F}$, $a \neq b$ with $\vec{y} + a\vec{x}, \vec{y} + b\vec{x} \in U_j, j \geq 2$ then

$$(a - b)\vec{x} = (\vec{y} + a\vec{x}) - (\vec{y} + b\vec{x}) \in U_j,$$

contrary to the choice of $\vec{x}$.

Conclusion: since $\mathbb{F}$ is infinite, L is infinite. But we have shewn that L can have at most one element in common with the U_j. This means that there are not enough U_j to go around to cover the whole of L. So V cannot be a finite union of proper subspaces.

4.2.10 Take $F = \mathbb{Z}_2, V = F \times F$. Then V has the four elements

$$\begin{bmatrix} \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{1} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{1} \end{bmatrix},$$

with the following subspaces

$$V_1 = \left\{ \begin{bmatrix} \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{0} \\ \overline{1} \end{bmatrix} \right\}, \quad V_2 = \left\{ \begin{bmatrix} \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{0} \end{bmatrix} \right\}, \quad V_3 = \left\{ \begin{bmatrix} \overline{0} \\ \overline{0} \end{bmatrix}, \begin{bmatrix} \overline{1} \\ \overline{1} \end{bmatrix} \right\}.$$

It is easy to verify that these subspaces satisfy the conditions of the problem.

4.3.1 If

$$a\begin{bmatrix}1\\0\\0\end{bmatrix}+b\begin{bmatrix}1\\1\\0\end{bmatrix}+c\begin{bmatrix}1\\1\\1\end{bmatrix}=\overrightarrow{0},$$

then

$$\begin{bmatrix}a+b+c\\b+c\\c\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}.$$

This clearly entails that $c = b = a = 0$, and so the family is free.

4.3.2 Assume

$$a\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+b\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}+c\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}+d\begin{bmatrix}1\\1\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}.$$

Then

$$\begin{aligned}a+b+c+d&=0,\\a+b-c+d&=0,\\a-b+c&=0,\\a-b-c+d&=0.\end{aligned}$$

Subtracting the second equation from the first, we deduce $2c = 0$, that is, $c = 0$. Subtracting the third equation from the fourth, we deduce $-2c + d = 0$ or $d = 0$. From the first and third equations, we then deduce $a + b = 0$ and $a - b = 0$, which entails $a = b = 0$. In conclusion, $a = b = c = d = 0$.

Now, put

$$x\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+y\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}+z\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}+w\begin{bmatrix}1\\1\\0\\1\end{bmatrix}=\begin{bmatrix}1\\2\\1\\1\end{bmatrix}.$$

Then

$$\begin{aligned}x+y+z+w&=1,\\x+y-z+w&=2,\\x-y+z&=1,\\x-y-z+w&=1.\end{aligned}$$

Solving as before, we find

$$2\begin{bmatrix}1\\1\\1\\1\end{bmatrix}+\frac{1}{2}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix}-\frac{1}{2}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix}-\begin{bmatrix}1\\1\\0\\1\end{bmatrix}=\begin{bmatrix}1\\2\\1\\1\end{bmatrix}.$$

4.3.5 We have

$$(\vec{v}_1 + \vec{v}_2) - (\vec{v}_2 + \vec{v}_3) + (\vec{v}_3 + \vec{v}_4) - (\vec{v}_4 + \vec{v}_1) = \vec{0},$$

a non-trivial linear combination of these vectors equalling the zero-vector.

4.3.7 Yes. Suppose that $a + b\sqrt{2} = 0$ is a non-trivial linear combination of 1 and $\sqrt{2}$ with rational numbers a and b. If one of a, b is different from 0 then so is the other. Hence

$$a + b\sqrt{2} = 0 \implies \sqrt{2} = -\frac{b}{a}.$$

The sinistral side of the equality $\sqrt{2} = -\frac{b}{a}$ is irrational whereas the dextral side is rational, a contradiction.

4.3.8 No. The representation $2 \cdot 1 + (-\sqrt{2})\sqrt{2} = 0$ is a non-trivial linear combination of 1 and $\sqrt{2}$.

4.3.9

1. Assume that

$$a + b\sqrt{2} + c\sqrt{3} = 0, \quad a, b, c, \in \mathbb{Q}, a^2 + b^2 + c^2 \neq 0.$$

If $ac \neq 0$, then

$$b\sqrt{2} = -a - c\sqrt{3} \Leftrightarrow 2b^2 = a^2 + 2ac\sqrt{3} + 3c^2 \Leftrightarrow \frac{2b^2 - a^2 - 3c^2}{2ac} = \sqrt{3}.$$

The dextral side of the last implication is irrational, whereas the sinistral side is rational. Thus it must be the case that $ac = 0$. If $a = 0$, $c \neq 0$ then

$$b\sqrt{2} + c\sqrt{3} = 0 \Leftrightarrow -\frac{b}{c} = \sqrt{\frac{3}{2}},$$

and again the dextral side is irrational and the sinistral side is rational. Thus if $a = 0$ then also $c = 0$. We can similarly prove that $c = 0$ entails $a = 0$. Thus we have

$$b\sqrt{2} = 0,$$

which means that $b = 0$. Therefore

$$a + b\sqrt{2} + c\sqrt{3} = 0, a, b, c, \in \mathbb{Q}, \Leftrightarrow a = b = c = 0.$$

This proves that $\{1, \sqrt{2}, \sqrt{3}\}$ are linearly independent over $\mathbb{Q}$.

2. Rationalising denominators,

$$\begin{aligned} \frac{1}{1-\sqrt{2}} + \frac{2}{\sqrt{12}-2} &= \frac{1+\sqrt{2}}{1-2} + \frac{2\sqrt{12}+4}{12-4} \\ &= -1 - \sqrt{2} + \frac{1}{2}\sqrt{3} + \frac{1}{2} \\ &= -\frac{1}{2} - \sqrt{2} + \frac{1}{2}\sqrt{3}. \end{aligned}$$

4.3.10 Assume that

$$ae^x + be^{2x} + ce^{3x} = 0.$$

Then

$$c = -ae^{-2x} - be^{-x}.$$

Letting $x \to +\infty$, we obtain $c = 0$. Thus

$$ae^x + be^{2x} = 0,$$

and so

$$b = -ae^{-x}.$$

Again, letting $x \to +\infty$, we obtain $b = 0$. This yields

$$ae^x = 0.$$

Since the exponential function never vanishes, we deduce that $a = 0$. Thus $a = b = c = 0$ and the family is linearly independent over $\mathbb{R}$.

4.3.11 This follows at once from the identity

$$\cos 2x = \cos^2 x - \sin^2 x,$$

which implies

$$\cos 2x - \cos^2 x + \sin^2 x = 0.$$

4.4.1 Given an arbitrary polynomial

$$p(x) = a + bx + cx^2 + dx^3,$$

we must shew that there are real numbers s, t, u, v such that

$$p(x) = s + t(1 + x) + u(1 + x)^2 + v(1 + x)^3.$$

In order to do this we find the Taylor expansion of p around $x = -1$. Letting $x = -1$ in this last equality,

$$s = p(-1) = a - b + c - d \in \mathbb{R}.$$

Now,

$$p'(x) = b + 2cx + 3dx^2 = t + 2u(1 + x) + 3v(1 + x)^2.$$

Letting $x = -1$ we find

$$t = p'(-1) = b - 2c + 3d \in \mathbb{R}.$$

Again,

$$p''(x) = 2c + 6dx = 2u + 6v(1 + x).$$

Letting $x = -1$ we find

$$u = p''(-1) = c - 3d \in \mathbb{R}.$$

Finally,

$$p'''(x) = 6d = 6v,$$

so we let $v = d \in \mathbb{R}$. In other words, we have

$$p(x) = a + bx + cx^2 + dx^3 = (a - b + c - d) + (b - 2c + 3d)(1 + x) + (c - 3d)(1 + x)^2 + d(1 + x)^3.$$

4.4.2 Assume contrariwise that

$$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Then we must have

$$a = 1,$$
$$b = 1,$$
$$-a - b = -1,$$

which is impossible. Thus $\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$ is not a linear combination of $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ and hence is not in **span** $\left(\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}\right)$.

4.4.3 It is

$$a \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + b \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} + c \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} a & c \\ -c & b \end{bmatrix},$$

i.e., this family spans the set of all skew-symmetric 2×2 matrices over $\mathbb{R}$.

4.5.1 We have

$$\begin{bmatrix} a \\ 2a - 3b \\ 5b \\ a + 2b \\ a \end{bmatrix} = a \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ -3 \\ 5 \\ 2 \\ 0 \end{bmatrix},$$

so clearly the family

$$\left\{\begin{bmatrix}1\\2\\0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\-3\\5\\2\\0\end{bmatrix}\right\}$$

spans the subspace. To shew that this is a linearly independent family, assume that

$$a\begin{bmatrix}1\\2\\0\\1\\1\end{bmatrix} + b\begin{bmatrix}0\\-3\\5\\2\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\\0\\0\end{bmatrix}.$$

Then it follows clearly that $a = b = 0$, and so this is a linearly independent family. Conclusion:

$$\left\{\begin{bmatrix}1\\2\\0\\1\\1\end{bmatrix}, \begin{bmatrix}0\\-3\\5\\2\\0\end{bmatrix}\right\}$$

is a basis for the subspace.

4.5.2 Suppose

$$\begin{aligned}\vec{0} &= a(\vec{v}_1 + \vec{v}_2) + b(\vec{v}_2 + \vec{v}_3) + c(\vec{v}_3 + \vec{v}_4) + d(\vec{v}_4 + \vec{v}_5) + f(\vec{v}_5 + \vec{v}_1) \\ &= (a + f)\vec{v}_1 + (a + b)\vec{v}_2 + (b + c)\vec{v}_3 + (c + d)\vec{v}_4 + (d + f)\vec{v}_5.\end{aligned}$$

Since $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_5\}$ are linearly independent, we have

$$a + f = 0,$$

$$a + b = 0$$

$$b + c = 0$$

$$c + d = 0$$

$$d + f = 0.$$

Solving we find $a = b = c = d = f = 0$, which means that the

$$\{\vec{v}_1 + \vec{v}_2, \vec{v}_2 + \vec{v}_3, \vec{v}_3 + \vec{v}_4, \vec{v}_4 + \vec{v}_5, \vec{v}_5 + \vec{v}_1\}$$

are linearly independent. Since the dimension of V is 5, and we have 5 linearly independent vectors, they must also be a basis for V.

4.5.3 The matrix of coefficients is already in echelon form. The dimension of the solution space is $n-1$ and the following vectors in $\mathbb{R}^{2n}$ form a basis for the solution space

$$a_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \\ \vdots \\ 0 \\ -1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad a_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ \vdots \\ 0 \\ -1 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \quad a_{n-1} = \begin{bmatrix} -1 \\ 0 \\ \ldots \\ 1 \\ -1 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

(The "second" -1 occurs on the n-th position. The 1's migrate from the 2nd and $n+1$-th position on a_1 to the $n-1$-th and $2n$-th position on a_{n-1}.)

4.5.4 Let $A^T = -A$ and $B^T = -B$ be skew symmetric $n \times n$ matrices. Then if $\lambda \in \mathbb{R}$ is a scalar, then

$$(A + \lambda B)^T = -(A + \lambda B),$$

so $A + \lambda B$ is also skew-symmetric, proving that V is a subspace. Now consider the set of

$$1 + 2 + \cdots + (n-1) = \frac{(n-1)n}{2}$$

matrices A_k, which are 0 everywhere except in the ij-th and ji-spot, where $1 \leq i < j \leq n$, $a_{ij} = 1 = -a_{ji}$ and $i+j = k$, $3 \leq k \leq 2n-1$. (In the case $n = 3$, they are

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},$$

for example.) It is clear that these matrices form a basis for V and hence V has dimension $\dfrac{(n-1)n}{2}$.

4.5.5 Take $(\vec{u}, \vec{v}) \in X^2$ and $\alpha \in \mathbb{R}$. Then

$$\vec{u} = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}, \quad b + 2c = 0, \quad \vec{v} = \begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix}, \quad b' + 2c' = 0.$$

We have

$$\vec{u}+\alpha\vec{v}=\begin{bmatrix} a+\alpha a' \\ b+\alpha b' \\ c+\alpha c' \\ d+\alpha d' \end{bmatrix},$$

and to demonstrate that $\vec{u}+\alpha\vec{v}\in X$ we need to shew that $(b+\alpha b')+2(c+\alpha c')=0$. But this is easy, as

$$(b+\alpha b')+2(c+\alpha c')=(b+2c)+\alpha(b'+2c')=0+\alpha 0=0.$$

Now

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}=\begin{bmatrix} a \\ -2c \\ c \\ d \end{bmatrix}=a\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}+c\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}+d\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

It is clear that

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},\begin{bmatrix} 0 \\ -2 \\ 1 \\ 0 \end{bmatrix},\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

are linearly independent and span X. They thus constitute a basis for X.

4.5.6 As a basis we may take the $\dfrac{n(n+1)}{2}$ matrices $\mathbf{E}_{ij}\in\mathbf{M}_n(\mathbb{F})$ for $1\leq i\leq j\leq n$.

4.5.7 $\mathbf{dim}\, X=2$, as basis one may take $\{\vec{v_1},\vec{v_2}\}$.

4.5.8 $\mathbf{dim}\, X=3$, as basis one may take $\{\vec{v_1},\vec{v_2},\vec{v_3}\}$.

4.5.9 $\mathbf{dim}\, X=3$, as basis one may take $\{\vec{v_1},\vec{v_2},\vec{v_3}\}$.

4.5.10 Let $\lambda\in\mathbb{R}$. Observe that

$$\begin{bmatrix} a & b & c \\ 0 & d & f \\ 0 & 0 & g \end{bmatrix}+\lambda\begin{bmatrix} a' & b' & c' \\ 0 & d' & f' \\ 0 & 0 & g' \end{bmatrix}=\begin{bmatrix} a+\lambda a' & b+\lambda b' & c+\lambda c' \\ 0 & d+\lambda d' & f+\lambda f' \\ 0 & 0 & g+\lambda g' \end{bmatrix}$$

and if $a+b+c=0,\quad a+d+g=0,\quad a'+b'+c'=0,\quad a'+d'+g'=0$, then

$$a+\lambda a'+b+\lambda b'+c+\lambda c'=(a+b+c)+\lambda(a'+b'+c')=0+\lambda 0=0,$$

and

$$a+\lambda a'+d+\lambda d'+g+\lambda g'=(a+d+g)+\lambda(a'+d'+g')=0+\lambda 0=0,$$

proving that $\mathbf{V}$ is a subspace.

Now, $a+b+c=0=a+d+g \implies a=-b-c,\ g=b+c-d$. Thus

$$\begin{bmatrix} a & b & c \\ 0 & d & f \\ 0 & 0 & g \end{bmatrix}=\begin{bmatrix} -b-c & b & c \\ 0 & d & f \\ 0 & 0 & b+c-d \end{bmatrix}=b\begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}+c\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}+d\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}+f\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

It is clear that these four matrices span V and are linearly independent. Hence, **dim** V = 4.

4.6.1 1. It is enough to prove that the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

is invertible. But an easy computation shews that

$$A^2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}^2 = 4\mathbf{I}_4,$$

whence the inverse sought is

$$A^{-1} = \frac{1}{4}A = \frac{1}{4}\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/4 & -1/4 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ 1/4 & -1/4 & -1/4 & 1/4 \end{bmatrix}.$$

2. Since the $\overrightarrow{a}_k$ are four linearly independent vectors in $\mathbb{R}^4$ and **dim** $\mathbb{R}^4 = 4$, they form a basis for $\mathbb{R}^4$. Now, we want to solve

$$A\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

and so

$$\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = A^{-1}\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/4 & -1/4 \\ 1/4 & -1/4 & 1/4 & -1/4 \\ 1/4 & -1/4 & -1/4 & 1/4 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5/4 \\ 1/4 \\ -1/4 \\ -1/4 \end{bmatrix}.$$

It follows that

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix} = \frac{5}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{4}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} - \frac{1}{4}\begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}.$$

The coordinates sought are

$$\left(\frac{5}{4}, \frac{1}{4}, -\frac{1}{4}, -\frac{1}{4}\right).$$

3. Since we have

$$\begin{bmatrix}1\\2\\1\\1\end{bmatrix} = \frac{5}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \frac{1}{4}\begin{bmatrix}1\\-1\\1\\-1\end{bmatrix} + \frac{1}{4}\begin{bmatrix}1\\1\\-1\\-1\end{bmatrix} - \frac{1}{4}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix},$$

the coordinates sought are

$$\left(\frac{5}{4}, -\frac{1}{4}, \frac{1}{4}, -\frac{1}{4}\right).$$

4.6.2 [1] $a = 1$, [2] $(A(a))^{-1} = \begin{bmatrix}\frac{1}{a-1} & 0 & 0 & -\frac{1}{a-1}\\ -1 & 1-a & -1 & a+1\\ -\frac{1}{a-1} & -1 & 0 & \frac{a}{a-1}\\ 1 & a & 1 & -a-1\end{bmatrix}$ [3]

$$\begin{bmatrix}0 & \frac{1}{a-1} & \frac{1}{a-1} & \frac{1}{a-1}\\ 0 & -a-1 & -a & -1\\ 0 & --\frac{a}{a-1} & -\frac{a}{a-1} & -\frac{1}{a-1}\\ 1 & 2+a & a+1 & 1\end{bmatrix}$$

5.1.1 Let $\alpha \in \mathbb{R}$. Then

$$\begin{aligned}
L\begin{bmatrix}x+\alpha a\\ y+\alpha b\\ z+\alpha c\end{bmatrix} &= \begin{bmatrix}(x+\alpha a)-(y+\alpha b)-(z+\alpha c)\\ (x+\alpha a)+(y+\alpha b)+(z+\alpha c)\\ z+\alpha c\end{bmatrix}\\
&= \begin{bmatrix}x-y-z\\ x+y+z\\ z\end{bmatrix} + \alpha\begin{bmatrix}a-b-c\\ a+b+c\\ c\end{bmatrix}\\
&= L\begin{bmatrix}x\\ y\\ z\end{bmatrix} + \alpha L\begin{bmatrix}a\\ b\\ c\end{bmatrix},
\end{aligned}$$

proving that L is a linear transformation.

5.1.2

$$\begin{aligned} L(H+\alpha H') &= -A^{-1}(H+\alpha H')A^{-1} \\ &= -A^{-1}HA^{-1} + \alpha(-A^{-1}H'A^{-1}) \\ &= L(H) + \alpha L(H'), \end{aligned}$$

proving that L is linear.

5.1.3 Let S be convex and let $\overrightarrow{a}, \overrightarrow{b} \in T(S)$. We must prove that $\forall \alpha \in [0;1]$, $(1-\alpha)\overrightarrow{a} + \alpha\overrightarrow{b} \in T(S)$. But since $\overrightarrow{a}, \overrightarrow{b}$ belong to $T(S)$, $\exists \overrightarrow{x} \in S, \overrightarrow{y} \in S$ with $T(\overrightarrow{x}) = \overrightarrow{a}, T(\overrightarrow{y}) = \overrightarrow{b}$. Since S is convex, $(1-\alpha)\overrightarrow{x} + \alpha\overrightarrow{y} \in S$. Thus

$$T((1-\alpha)\overrightarrow{x} + \alpha\overrightarrow{y}) \in T(S),$$

which means that

$$(1-\alpha)T(\overrightarrow{x}) + \alpha T(\overrightarrow{y}) \in T(S),$$

that is,

$$(1-\alpha)\overrightarrow{a} + \alpha\overrightarrow{b} \in T(S),$$

as we wished to show.

5.2.1 Assume $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbf{ker}\,(L)$. Then

$$L\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

that is

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

This implies that $x - y = 0$ and $x + y = 0$, and so $x = y = z = 0$. This means that

$$\mathbf{ker}\,(L) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\},$$

and L is injective.

By the Dimension Theorem 244, $\mathbf{dim}\,\mathbf{Im}\,(L) = \mathbf{dim}\,V - \mathbf{dim}\,\mathbf{ker}\,(L) = 3 - 0 = 3$, which means that

$$\mathbf{Im}\,(L) = \mathbb{R}^3$$

and L is surjective.

5.2.2

1. If a is any scalar,

$$L\left(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} + a\begin{bmatrix} x' \\ y' \\ z' \\ w' \end{bmatrix}\right) = L\begin{bmatrix} x + ax' \\ y + ay' \\ z + az' \\ w + aw' \end{bmatrix} = \begin{bmatrix} (x+ax') + (y+ay') \\ (x+ax') - (y+ay') \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix} + a\begin{bmatrix} x'+y' \\ x'-y' \end{bmatrix} = L\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} + aL\begin{bmatrix} x' \\ y' \\ z' \\ w' \end{bmatrix},$$

whence L is linear.

2. We have,

$$L\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} x+y \\ x-y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \implies x = y, x = -y \implies x = y = 0 \implies \mathbf{ker}(L) = \left\{ \begin{bmatrix} 0 \\ 0 \\ z \\ w \end{bmatrix} : z \in \mathbb{R}, w \in \mathbb{R} \right\}.$$

Thus $\mathbf{dim}\,\mathbf{ker}(L) = 2$. In particular, the transformation is not injective.

3. From the previous part, $\mathbf{dim}\,\mathbf{Im}(L) = 4 - 2 = 2$. Since $\mathbf{Im}(L) \subseteq \mathbb{R}^2$ and $\mathbf{dim}\,\mathbf{Im}(L) = 2$, we must have $\mathbf{Im}(L) = \mathbb{R}^2$. In particular, the transformation is surjective.

5.2.3 Assume that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbf{ker}(T)$,

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = (a-b)\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then

$$\begin{aligned} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} &= T\begin{bmatrix} a \\ b \\ c \end{bmatrix} \\ &= (a-b)T\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + bT\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + cT\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= (a-b)\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + b\begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \end{bmatrix} + c\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \\ &= \begin{bmatrix} a+b+c \\ -b-c \\ -a+b+c \\ 0 \end{bmatrix}. \end{aligned}$$

It follows that $a = 0$ and $b = -c$. Thus

$$\mathbf{ker}\,(T) = \left\{ c \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} : c \in \mathbb{R} \right\},$$

and so $\mathbf{dim}\,\mathbf{ker}\,(T) = 1$.

By the Dimension Theorem 244,

$$\mathbf{dim}\,\mathbf{Im}\,(T) = \mathbf{dim}\,V - \mathbf{dim}\,\mathbf{ker}\,(T) = 3 - 1 = 2.$$

We readily see that

$$\begin{bmatrix} 2 \\ -1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix},$$

and so

$$\mathbf{Im}\,(T) = \mathbf{span}\left(\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \right).$$

5.2.4 Assume that

$$L \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ x + 2y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then $x = -2y$ and so

$$\begin{bmatrix} x \\ y \end{bmatrix} = y \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

This means that $\mathbf{dim}\,\mathbf{ker}\,(L) = 1$ and $\mathbf{ker}\,(L)$ is the line through the origin and $(-2, 1)$. Observe that L is not injective.

By the Dimension Theorem 244, $\mathbf{dim}\,\mathbf{Im}\,(L) = \mathbf{dim}\,V - \mathbf{dim}\,\mathbf{ker}\,(L) = 2 - 1 = 1$. Assume that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbf{Im}\,(L)$.

Then $\exists (x, y) \in \mathbb{R}^2$ such that

$$L \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x + 2y \\ x + 2y \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

This means that

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x+2y \\ x+2y \\ 0 \end{bmatrix} = (x+2y)\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Observe that L is not surjective.

5.2.5 Assume that

$$L\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ x+y \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then $x+y=0=x-y$, that is, $x=y=0$, meaning that

$$\mathbf{ker}(L) = \left\{\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\},$$

and so L is injective.

By the Dimension Theorem 244, $\mathbf{dim}\,\mathbf{Im}(L) = \mathbf{dim}\,V - \mathbf{dim}\,\mathbf{ker}(L) = 2 - 0 = 2$. Assume that $\begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbf{Im}(L)$.

Then $\exists(x,y) \in \mathbb{R}^2$ such that

$$L\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x-y \\ x+y \\ 0 \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

This means that

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x-y \\ x+y \\ 0 \end{bmatrix} = x\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Since

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

are linearly independent, they span a subspace of dimension 2 in $\mathbb{R}^3$, that is, a plane containing the origin. Observe that L is not surjective.

5.2.6 Assume that

$$L\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y-z \\ y-2z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Then $y = 2z; x = y + z = 3z$. This means that $\mathbf{ker}(L) = \left\{ z \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} : z \in \mathbb{R} \right\}$. Hence $\mathbf{dim}\,\mathbf{ker}(L) = 1$, and so L is not injective.

Now, if

$$L \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y - z \\ y - 2z \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Then

$$\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x - y - z \\ y - 2z \end{bmatrix} = x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -1 \\ 1 \end{bmatrix} + z \begin{bmatrix} -1 \\ -2 \end{bmatrix}.$$

Now,

$$-3 \begin{bmatrix} 1 \\ 0 \end{bmatrix} - 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

are linearly independent. Since $\mathbf{dim}\,\mathbf{Im}(L) = 2$, we have $\mathbf{Im}(L) = \mathbb{R}^2$, and so L is surjective.

5.2.7 Assume that

$$0 = \mathbf{tr}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + d.$$

Then $a = -d$ and so,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} -d & b \\ c & d \end{bmatrix} = d \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix},$$

and so $\mathbf{dim}\,\mathbf{ker}(L) = 3$. Thus L is not injective. L is surjective, however. For if $\alpha \in \mathbb{R}$, then

$$\alpha = \mathbf{tr}\left(\begin{bmatrix} \alpha & 0 \\ 0 & 0 \end{bmatrix}\right).$$

5.2.8 1. Let $(A, B)^2 \in \mathbf{M}_{2\times 2}(\mathbb{R})$, $\alpha \in \mathbb{R}$. Then

$$\begin{aligned} L(A + \alpha B) &= (A + \alpha B)^T + (A + \alpha B) \\ &= A^T + B^T + A + \alpha B \\ &= A^T + A + \alpha B^T + \alpha B \\ &= L(A) + \alpha L(B), \end{aligned}$$

proving that L is linear.

2. Assume that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbf{ker}\,(L)$. Then

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = L(A) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \begin{bmatrix} 2a & b+c \\ b+c & 2d \end{bmatrix},$$

whence $a = d = 0$ and $b = -c$. Hence

$$\mathbf{ker}\,(L) = \mathbf{span}\left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right),$$

and so $\mathbf{dim}\,\mathbf{ker}\,(L) = 1$.

3. By the Dimension Theorem, $\mathbf{dim}\,\mathbf{Im}\,(L) = 4 - 1 = 3$. As above,

$$\begin{aligned} L(A) &= \begin{bmatrix} 2a & b+c \\ b+c & 2d \end{bmatrix} \\ &= a\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} + (b+c)\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \end{aligned}$$

from where

$$\mathbf{Im}\,(L) = \mathbf{span}\left(\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}\right).$$

5.2.9 ❶ Observe that

$$(I - T)^2 = I - 2T + T^2 = I - 2T + T = I - T,$$

proving the result.

❷ The inverse is $I - \frac{1}{2}T$, for

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

proving the claim.

❸ We have

$$\begin{aligned} \overrightarrow{x} \in \mathbf{ker}\,(T) &\iff \overrightarrow{x} - T(\overrightarrow{x}) \in \mathbf{ker}\,(T) \\ &\iff I(\overrightarrow{x}) - T(\overrightarrow{x}) \in \mathbf{ker}\,(T) \\ &\iff (I - T)(\overrightarrow{x}) \in \mathbf{ker}\,(T) \\ &\iff \overrightarrow{x} \in \mathbf{Im}\,(I - T). \end{aligned}$$

5.3.1 Observe that

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = d\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + (2a - c - b)\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + (-d - 2a + 2c + b)\begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + (-a + c)\begin{bmatrix} -1 \\ -2 \\ 0 \\ 0 \end{bmatrix}.$$

Hence

$$T\begin{bmatrix}a\\b\\c\\d\end{bmatrix} = dT\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + (2a-c-b)T\begin{bmatrix}1\\0\\1\\0\end{bmatrix} + (-d-2a+2c+b)T\begin{bmatrix}1\\1\\1\\0\end{bmatrix} + (-a+c)T\begin{bmatrix}-1\\-2\\0\\0\end{bmatrix}$$

$$= d\begin{bmatrix}0\\0\\1\end{bmatrix} + (2a-c-b)\begin{bmatrix}1\\1\\-1\end{bmatrix} + (-d-2a+2c+b)\begin{bmatrix}0\\0\\-1\end{bmatrix} + (-a+c)\begin{bmatrix}1\\1\\1\end{bmatrix}$$

$$= \begin{bmatrix}a-b\\a-b\\-a+2d\end{bmatrix}.$$

This gives

$$T\begin{bmatrix}1\\0\\0\\0\end{bmatrix} = \begin{bmatrix}1\\1\\-1\end{bmatrix}, \quad T\begin{bmatrix}0\\1\\0\\0\end{bmatrix} = \begin{bmatrix}-1\\-1\\0\end{bmatrix}, \quad T\begin{bmatrix}0\\0\\1\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}, \quad T\begin{bmatrix}0\\0\\0\\1\end{bmatrix} = \begin{bmatrix}0\\0\\2\end{bmatrix}.$$

The required matrix is therefore

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

This matrix has rank 2, and so $\mathbf{dim}\,\mathbf{Im}\,(T) = 2$. We can use $\left\{\begin{bmatrix}1\\1\\-1\end{bmatrix}, \begin{bmatrix}-1\\-1\\0\end{bmatrix}\right\}$ as a basis for $\mathbf{Im}\,(T)$. Thus by the dimension theorem $\mathbf{dim}\,\mathbf{ker}\,(T) = 2$. If $\begin{bmatrix}0\\0\\0\end{bmatrix} = T\begin{bmatrix}a\\b\\c\\d\end{bmatrix} = \begin{bmatrix}a-b\\a-b\\-a+2d\end{bmatrix}$, Hence the vectors in $\mathbf{ker}\,(T)$ have the form $\begin{bmatrix}2d\\2d\\c\\d\end{bmatrix}$ and hence we may take $\left\{\begin{bmatrix}2\\2\\0\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right\}$ as a basis for $\mathbf{ker}\,(T)$.

5.3.2 1. Since the image of T is the plane $x+y+z=0$, we must have

$$a+0+1=0 \implies a=-1,$$

$$3+b-5=0 \implies b=2,$$

$$-1+2+c=0 \implies c=-1.$$

2. Observe that $\begin{bmatrix}1\\1\\1\end{bmatrix} \in \mathbf{ker}\,(T)$ and so

$$T\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Thus

$$T\begin{bmatrix}1\\0\\0\end{bmatrix} = T\begin{bmatrix}2\\1\\1\end{bmatrix} - T\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}3\\2\\-5\end{bmatrix},$$

$$T\begin{bmatrix}0\\1\\0\end{bmatrix} = T\begin{bmatrix}1\\2\\1\end{bmatrix} - T\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-1\\2\\-1\end{bmatrix},$$

$$T\begin{bmatrix}0\\0\\1\end{bmatrix} = T\begin{bmatrix}1\\1\\2\end{bmatrix} - T\begin{bmatrix}1\\1\\1\end{bmatrix} = \begin{bmatrix}-1\\0\\1\end{bmatrix}.$$

The required matrix is therefore

$$\begin{bmatrix}3 & -1 & -1\\2 & 2 & 0\\-5 & -1 & 1\end{bmatrix}.$$

5.3.3 1. Let $\alpha \in \mathbb{R}$. We have

$$\begin{aligned} T\left(\begin{bmatrix} x \\ y \end{bmatrix} + \alpha \begin{bmatrix} u \\ v \end{bmatrix}\right) &= T\left(\begin{bmatrix} x + \alpha u \\ y + \alpha v \end{bmatrix}\right) \\ &= \begin{bmatrix} x + \alpha u + y + \alpha v \\ x + \alpha u - y - \alpha v \\ 2(x + \alpha u) + 3(y + \alpha v) \end{bmatrix} \\ &= \begin{bmatrix} x + y \\ x - y \\ 2x + 3y \end{bmatrix} + \alpha \begin{bmatrix} u + v \\ u - v \\ 2u + 3v \end{bmatrix} \\ &= T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) + \alpha T\left(\begin{bmatrix} u \\ v \end{bmatrix}\right), \end{aligned}$$

proving that T is linear.

2. We have

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff \begin{bmatrix} x + y \\ x - y \\ 2x + 3y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff x = y = 0,$$

$\mathbf{dim}\,\mathbf{ker}\,(T) = 0$, and whence T is injective.

3. By the Dimension Theorem, $\mathbf{dim}\,\mathbf{Im}\,(T) = 2 - 0 = 2$. Now, since

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x + y \\ x - y \\ 2x + 3y \end{bmatrix} = x \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix},$$

whence

$$\mathbf{Im}\,(T) = \mathbf{span}\left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix}\right).$$

4. We have

$$T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ -1 \\ 8 \end{bmatrix} = \frac{11}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{5}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{13}{2}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 11/2 \\ -5/2 \\ -13/2 \end{bmatrix}_{\mathscr{B}},$$

and

$$T\left(\begin{bmatrix} 1 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ -2 \\ 11 \end{bmatrix} = \frac{15}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{19}{2}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 15/2 \\ -7/2 \\ -19/2 \end{bmatrix}_{\mathscr{B}}.$$

The required matrix is

$$\begin{bmatrix} 11/2 & 15/2 \\ -5/2 & -7/2 \\ -13/2 & -19/2 \end{bmatrix}_{\mathscr{B}}.$$

5.3.4 The matrix will be a 2×3 matrix. In each case, we find the action of **L** on the basis elements of $\mathbb{R}^3$ and express the result in the given basis for $\mathbb{R}^3$.

1. We have

$$\mathbf{L}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{L}\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \mathbf{L}\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

The required matrix is

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

2. We have

$$\mathbf{L}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \mathbf{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \mathbf{L}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

The required matrix is

$$\begin{bmatrix} 1 & 3 & 3 \\ 3 & 3 & 2 \end{bmatrix}.$$

3. We have

$$\mathbf{L}\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = -2\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}_{\mathscr{A}},$$

$$\mathbf{L}\left(\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 3 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}_{\mathscr{A}},$$

$$\mathbf{L}\left(\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 3 \\ 2 \end{bmatrix} = 1\begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}_{\mathscr{A}}.$$

The required matrix is

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

5.3.5 Observe that $\begin{bmatrix} 2 \\ 3 \end{bmatrix} \in \mathbf{Im}\,(T) = \mathbf{ker}\,(T)$ and so

$$T\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Now

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = T\left(3\begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) = 3T\begin{bmatrix} 1 \\ 1 \end{bmatrix} - T\begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 6 \\ 9 \end{bmatrix},$$

and

$$T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = T\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix} - 2\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = T\begin{bmatrix} 2 \\ 3 \end{bmatrix} - 2T\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -4 \\ -6 \end{bmatrix}.$$

The required matrix is thus

$$\begin{bmatrix} 6 & -4 \\ 9 & -6 \end{bmatrix}.$$

5.3.6 The matrix will be a 1×4 matrix. We have

$$\mathbf{tr}\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) = 1,$$

$$\mathbf{tr}\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) = 0,$$

$$\mathbf{tr}\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) = 0,$$

$$\mathbf{tr}\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) = 1.$$

Thus

$$M_L = (1 \;\; 0 \;\; 0 \;\; 1).$$

5.3.7 First observe that $\mathbf{ker}\,(B) \subseteq \mathbf{ker}\,(AB)$ since $\forall X \in \mathbf{M}_{q\times 1}(\mathbb{R})$,

$$BX = 0 \implies (AB)X = A(BX) = 0.$$

Now

$$\begin{aligned} \mathbf{dim}\,\mathbf{ker}\,(B) &= q - \mathbf{dim}\,\mathbf{Im}\,(B) \\ &= q - \mathbf{rank}\,(B) \\ &= q - \mathbf{rank}\,(AB) \\ &= q - \mathbf{dim}\,\mathbf{Im}\,(AB) \\ &= \mathbf{dim}\,\mathbf{ker}\,(AB). \end{aligned}$$

Thus $\mathbf{ker}\,(B) = \mathbf{ker}\,(AB)$. Similarly, we can demonstrate that $\mathbf{ker}\,(ABC) = \mathbf{ker}\,(BC)$. Thus

$$\begin{aligned}\mathbf{rank}\,(ABC) &= \mathbf{dim}\,\mathbf{Im}\,(ABC)\\ &= r - \mathbf{dim}\,\mathbf{ker}\,(ABC)\\ &= r - \mathbf{dim}\,\mathbf{ker}\,(BC)\\ &= \mathbf{dim}\,\mathbf{Im}\,(BC)\\ &= \mathbf{rank}\,(BC)\,.\end{aligned}$$

6.2.1 This is clearly $(1\ 2\ 3\ 4)(6\ 8\ 7)$ of order $\mathbf{lcm}(4,3) = 12$.

6.3.1 Multiplying the first column of the given matrix by a, its second column by b, and its third column by c, we obtain

$$abc\Omega = \begin{bmatrix} abc & abc & abc \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{bmatrix}.$$

We may factor out abc from the first row of this last matrix thereby obtaining

$$abc\Omega = abc\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{bmatrix}.$$

Upon dividing by abc,

$$\Omega = \mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{bmatrix}.$$

6.3.2 Performing $R_1 + R_2 + R_3 \to R_1$ we have

$$\begin{aligned}\Omega = \mathbf{det}&\begin{bmatrix} a-b-c & 2a & 2a \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{bmatrix}\\ = \mathbf{det}&\begin{bmatrix} a+b+c & a+b+c & a+b+c \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{bmatrix}.\end{aligned}$$

Factorising $(a+b+c)$ from the first row of this last determinant, we have

$$\Omega = (a+b+c)\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ 2b & b-c-a & 2b \\ 2c & 2c & c-a-b \end{bmatrix}.$$

Performing $C_2 - C_1 \to C_2$ and $C_3 - C_1 \to C_3$,

$$\Omega = (a+b+c)\,\mathbf{det}\begin{bmatrix} 1 & 0 & 0 \\ 2b & -b-c-a & 0 \\ 2c & 0 & -c-a-b \end{bmatrix}.$$

This last matrix is triangular, hence

$$\Omega = (a+b+c)(-b-c-a)(-c-a-b) = (a+b+c)^3,$$

as wanted.

6.3.3 $\mathbf{det}\,A_1 = \mathbf{det}\,A = -540$ by multilinearity. $\mathbf{det}\,A_2 = -\mathbf{det}\,A_1 = 540$ by alternancy. $\mathbf{det}\,A_3 = 3\,\mathbf{det}\,A_2 = 1620$ by both multilinearity and homogeneity from one column. $\mathbf{det}\,A_4 = \mathbf{det}\,A_3 = 1620$ by multilinearity, and $\mathbf{det}\,A_5 = 2\,\mathbf{det}\,A_4 = 3240$ by homogeneity from one column.

6.3.5 From the given data, $\mathbf{det}\,B = -2$. Hence

$$\mathbf{det}\,ABC = (\mathbf{det}\,A)(\mathbf{det}\,B)(\mathbf{det}\,C) = -12,$$

$$\mathbf{det}\,5AC = 5^3\,\mathbf{det}\,AC = (125)(\mathbf{det}\,A)(\mathbf{det}\,C) = 750,$$

$$(\mathbf{det}\,A^3B^{-3}C^{-1}) = \frac{(\mathbf{det}\,A)^3}{(\mathbf{det}\,B)^3(\mathbf{det}\,C)} = -\frac{27}{16}.$$

6.3.6 Pick $\lambda \in \mathbb{R} \setminus \{0, a_{11}, a_{22}, \ldots, a_{nn}\}$. Put

$$X = \begin{bmatrix} a_{11}-\lambda & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22}-\lambda & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33}-\lambda & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn}-\lambda \end{bmatrix}$$

and

$$Y = \begin{bmatrix} \lambda & a_{12} & a_{13} & \vdots & a_{1n} \\ 0 & \lambda & a_{23} & \vdots & a_{2n} \\ 0 & 0 & \lambda & \vdots & a_{3n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \vdots & \lambda \end{bmatrix}$$

Clearly $A = X + Y$, $\mathbf{det}\,X = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) \neq 0$, and $\mathbf{det}\,Y = \lambda^n \neq 0$. This completes the proof.

6.3.7 No.

6.4.1 We have

$$\begin{aligned} \mathbf{det}\,A &= 2(-1)^{1+2}\,\mathbf{det}\begin{bmatrix} 4 & 6 \\ 7 & 9 \end{bmatrix} + 5(-1)^{2+2}\,\mathbf{det}\begin{bmatrix} 1 & 3 \\ 7 & 9 \end{bmatrix} + 8(-1)^{2+3}\,\mathbf{det}\begin{bmatrix} 1 & 3 \\ 4 & 6 \end{bmatrix} \\ &= -2(36-42) + 5(9-21) - 8(6-12) = 0. \end{aligned}$$

6.4.2 Simply expand along the first row

$$a\,\mathbf{det}\begin{bmatrix} a & b \\ c & a \end{bmatrix} - b\,\mathbf{det}\begin{bmatrix} c & b \\ b & a \end{bmatrix} + c\,\mathbf{det}\begin{bmatrix} c & a \\ b & c \end{bmatrix} = a(a^2 - bc) - b(ca - b^2) + c(c^2 - ab) = a^3 + b^3 + c^3 - 3abc.$$

6.4.3 Since the second column has three 0's, it is advantageous to expand along it, and thus we are reduced to calculate

$$-3(-1)^{3+2}\,\mathbf{det}\begin{bmatrix} 1 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Expanding this last determinant along the second column, the original determinant is thus

$$-3(-1)^{3+2}(-1)(-1)^{1+2}\,\mathbf{det}\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} = -3(-1)(-1)(-1)(1) = 3.$$

6.4.5 Expanding along the first column,

$$\begin{aligned} 0 &= \mathbf{det}\begin{bmatrix} 1 & 1 & 1 & 1 \\ x & a & 0 & 0 \\ x & 0 & b & 0 \\ x & 0 & 0 & c \end{bmatrix} \\ &= \mathbf{det}\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} - x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \\ &\quad + x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & 0 & 0 \\ 0 & 0 & c \end{bmatrix} - x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & 0 & 0 \\ 0 & b & 0 \end{bmatrix} \\ &= xabc - xbc + x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & 0 & 0 \\ 0 & 0 & c \end{bmatrix} - x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & 0 & 0 \\ 0 & b & 0 \end{bmatrix}. \end{aligned}$$

Expanding these last two determinants along the third row,

$$\begin{aligned} 0 &= abc - xbc + x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & 0 & 0 \\ 0 & 0 & c \end{bmatrix} - x\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 \\ a & a & 0 \\ 0 & b & 0 \end{bmatrix} \\ &= abc - xbc + xc\,\mathbf{det}\begin{bmatrix} 1 & 1 \\ a & 0 \end{bmatrix} + xb\,\mathbf{det}\begin{bmatrix} 1 & 1 \\ a & 0 \end{bmatrix} \\ &= abc - xbc - xca - xab. \end{aligned}$$

It follows that

$$abc = x(bc + ab + ca),$$

whence

$$\frac{1}{x} = \frac{bc + ab + ca}{abc} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c},$$

as wanted.

6.4.7 Expanding along the first row the determinant equals

$$\begin{aligned} -a\,\mathbf{det}\begin{bmatrix} a & b & 0 \\ 0 & 0 & b \\ 1 & 1 & 1 \end{bmatrix} + b\,\mathbf{det}\begin{bmatrix} a & 0 & 0 \\ 0 & a & b \\ 1 & 1 & 1 \end{bmatrix} &= ab\,\mathbf{det}\begin{bmatrix} a & b \\ 1 & 1 \end{bmatrix} + ab\,\mathbf{det}\begin{bmatrix} a & b \\ 1 & 1 \end{bmatrix} \\ &= 2ab(a - b), \end{aligned}$$

as wanted.

6.4.8 Expanding along the first row, the determinant equals

$$a\,\mathbf{det}\begin{bmatrix} a & 0 & b \\ 0 & d & 0 \\ c & 0 & d \end{bmatrix} + b\,\mathbf{det}\begin{bmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & c & d \end{bmatrix}.$$

Expanding the resulting two determinants along the second row, we obtain

$$ad\,\mathbf{det}\begin{bmatrix} a & b \\ c & d \end{bmatrix} + b(-c)\,\mathbf{det}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad(ad - bc) - bc(ad - bc) = (ad - bc)^2,$$

as wanted.

6.4.9 For $n = 1$ we have $\mathbf{det}(1) = 1 = (-1)^{1+1}$. For $n = 2$ we have

$$\mathbf{det}\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = -1 = (-1)^{2+1}.$$

Assume that the result is true for $n-1$. Expanding the determinant along the first column

$$\begin{aligned}
\mathbf{det}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} &= 1\,\mathbf{det}\begin{bmatrix} 0 & 0 & \vdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \cdots & \vdots & \vdots & \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \\
&\quad -1\,\mathbf{det}\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & \cdots & \vdots & 0 \\ 0 & 1 & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots & \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \\
&= 1(0) - (1)(-1)^{n} \\
&= (-1)^{n+1},
\end{aligned}$$

giving the result.

6.4.10 Perform $C_k - C_1 \to C_k$ for $k \in [2; n]$. Observe that these operations do not affect the value of the determinant. Then

$$\mathbf{det}\, A = \mathbf{det}\begin{bmatrix} 1 & n-1 & n-1 & n-1 & \cdots & n-1 \\ n & 2-n & 0 & 0 & \vdots & 0 \\ n & 0 & 3-n & 0 & \cdots & 0 \\ n & 0 & 0 & 4-n & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \\ n & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Expand this last determinant along the n-th row, obtaining,

$$\begin{aligned}
\det A &= (-1)^{1+n} n \det \begin{bmatrix} n-1 & n-1 & n-1 & \cdots & n-1 & n-1 \\ 2-n & 0 & 0 & \vdots & 0 & 0 \\ 0 & 3-n & 0 & \cdots & 0 & 0 \\ 0 & 0 & 4-n & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix} \\
&= (-1)^{1+n} n(n-1)(2-n)(3-n) \\
&\quad \cdots (-2)(-1) \det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \\
&= -(n!) \det \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 0 & 0 & \vdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \\
&= -(n!)(-1)^n \\
&= (-1)^{n+1} n!,
\end{aligned}$$

upon using the result of problem 6.4.9.

6.4.11 Recall that $\binom{n}{k} = \binom{n}{n-k}$,

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n$$

and

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0, \quad \text{if } n > 0.$$

Assume that n is odd. Observe that then there are $n+1$ (an even number) of columns and that on the same row, $\binom{n}{k}$ is on a column of opposite parity to that of $\binom{n}{n-k}$. By performing $C_1 - C_2 + C_3 - C_4 + \cdots + C_n - C_{n+1} \to C_1$, the first column becomes all 0's, whence the determinant if 0 if n is odd.

6.4.15 I will prove that

$$\det\begin{bmatrix}(b+c)^2 & ab & ac\\ ab & (a+c)^2 & bc\\ ac & bc & (a+b)^2\end{bmatrix} = 2abc(a+b+c)^3.$$

Using permissible row and column operations,

$$\begin{aligned}\det\begin{bmatrix}(b+c)^2 & ab & ac\\ ab & (a+c)^2 & bc\\ ac & bc & (a+b)^2\end{bmatrix} &= \det\begin{bmatrix}b^2+2bc+c^2 & ab & ac\\ ab & a^2+2ca+c^2 & bc\\ ac & bc & a^2+2ab+b^2\end{bmatrix}\\ &\overset{C_1+C_2+C_3\to C_1}{=} \det\begin{bmatrix}b^2+2bc+c^2+ab+ac & ab & ac\\ ab+a^2+2ca+c^2+bc & a^2+2ca+c^2 & bc\\ ac+bc+a^2+2ab+b^2 & bc & a^2+2ab+b^2\end{bmatrix}\\ &= \det\begin{bmatrix}(b+c)(a+b+c) & ab & ac\\ (a+c)(a+b+c) & a^2+2ca+c^2 & bc\\ (a+b)(a+b+c) & bc & a^2+2ab+b^2\end{bmatrix}\end{aligned}$$

Pulling out a factor, the above equals

$$(a+b+c)\det\begin{bmatrix}b+c & ab & ac\\ a+c & a^2+2ca+c^2 & bc\\ a+b & bc & a^2+2ab+b^2\end{bmatrix}$$

and performing $R_1+R_2+R_3\to R_1$, this is

$$(a+b+c)\det\begin{bmatrix}2a+2b+2c & ab+a^2+2ca+c^2+bc & ac+bc+a^2+2ab+b^2\\ a+c & a^2+2ca+c^2 & bc\\ a+b & bc & a^2+2ab+b^2\end{bmatrix}$$

Factoring this is

$$(a+b+c)\det\begin{bmatrix}2(a+b+c) & (a+c)(a+b+c) & (a+b)(a+b+c)\\ a+c & a^2+2ca+c^2 & bc\\ a+b & bc & a^2+2ab+b^2\end{bmatrix},$$

which in turn is

$$(a+b+c)^2\det\begin{bmatrix}2 & a+c & a+b\\ a+c & a^2+2ca+c^2 & bc\\ a+b & bc & a^2+2ab+b^2\end{bmatrix}$$

Performing $\mathbf{C}_2 - (a+c)\mathbf{C}_1 \to \mathbf{C}_2$ and $\mathbf{C}_3 - (a+b)\mathbf{C}_1 \to \mathbf{C}_3$ we obtain

$$(a+b+c)^2 \, \mathbf{det} \begin{bmatrix} 2 & -a-c & -a-b \\ a+c & 0 & -a^2-ab-ac \\ a+b & -a^2-ab-ac & 0 \end{bmatrix}$$

This last matrix we will expand by the second column, obtaining that the original determinant is thus

$$(a+b+c)^2 \left((a+c)\, \mathbf{det} \begin{bmatrix} a+c & -a^2-ab-ac \\ a+b & 0 \end{bmatrix} + (a^2+ab+ac)\, \mathbf{det} \begin{bmatrix} 2 & -a-b \\ a+c & -a^2-ab-ac \end{bmatrix} \right)$$

This simplifies to

$$\begin{aligned} (a+b+c)^2 \big((a+c)(a+b)(a^2+ab+ac) & \\ +(a^2+ab+ac)(-a^2-ab-ac+bc) \big) &= a(a+b+c)^3((a+c)(a+b)-a^2-ab-ac+bc) \\ &= 2abc(a+b+c)^3, \end{aligned}$$

as claimed.

6.4.16 We have

$$
\begin{aligned}
\mathbf{det}\begin{bmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{bmatrix}
&\overset{R_1+R_2+R_3+R_4 \to R_1}{\rightsquigarrow \atop =} \mathbf{det}\begin{bmatrix} a+b+c+d & a+b+c+d & a+b+c+d & a+b+c+d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{bmatrix} \\
&= (a+b+c+d)\,\mathbf{det}\begin{bmatrix} 1 & 1 & 1 & 1 \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{bmatrix} \\
&\overset{C_4-C_3+C_2-C_1 \to C_4}{\rightsquigarrow \atop =} (a+b+c+d)\begin{vmatrix} 1 & 1 & 1 & 0 \\ d & a & b & c-b+a-d \\ c & d & a & b-a+d-c \\ b & c & d & a-d+c-b \end{vmatrix} \\
&= (a+b+c+d)(a-b+c-d)\begin{vmatrix} 1 & 1 & 1 & 0 \\ d & a & b & 1 \\ c & d & a & -1 \\ b & c & d & 1 \end{vmatrix} \\
&\overset{R_2+R_3 \to R_2,\ R_4+R_3 \to R_4}{\rightsquigarrow \atop =} (a+b+c+d)(a-b+c-d)\begin{vmatrix} 1 & 1 & 1 & 0 \\ d+c & a+d & b+a & 0 \\ c & d & a & -1 \\ b+c & c+d & a+d & 0 \end{vmatrix} \\
&= (a+b+c+d)(a-b+c-d)\begin{vmatrix} 1 & 1 & 1 \\ d+c & a+d & b+a \\ b+c & c+d & a+d \end{vmatrix} \\
&\overset{C_1-C_3 \to C_1,\ C_2-C_3 \to C_2}{\rightsquigarrow \atop =} (a+b+c+d)(a-b+c-d)\begin{vmatrix} 0 & 0 & 1 \\ d+c-b-a & d-b & b+a \\ b+c-a-d & c-a & a+d \end{vmatrix} \\
&= (a+b+c+d)(a-b+c-d)\begin{vmatrix} d+c-b-a & d-b \\ b+c-a-d & c-a \end{vmatrix} \\
&= (a+b+c+d)(a-b+c-d)(d+c-b-a)(c-a)-(d-b)(b+c-a-d) \\
&= (a+b+c+d)(a-b+c-d) \\
&\qquad ((c-a)(c-a)+(c-a)(d-b)-(d-b)(c-a)-(d-b)(b-d)) \\
&= (a+b+c+d)(a-b+c-d)((a-c)^2+(b-d)^2).
\end{aligned}
$$

☞ *Since*

$$(a-c)^2 + (b-d)^2 = (a-c+i(b-d))(a-c-i(b-d)),$$

the above determinant is then

$$(a+b+c+d)(a-b+c-d)(a+ib-c-id)(a-ib-c+id).$$

Generalisations of this determinant are possible using roots of unity.

7.2.1 We have

$$\mathbf{det}(\lambda \mathbf{I}_2 - A) = \mathbf{det}\begin{bmatrix} \lambda-1 & 1 \\ 1 & \lambda-1 \end{bmatrix} = (\lambda-1)^2 - 1 = \lambda(\lambda-2),$$

whence the eigenvalues are 0 and 2. For $\lambda = 0$ we have

$$0\mathbf{I}_2 - A = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}.$$

This has row-echelon form

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

If

$$\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

then $a = b$. Thus

$$\begin{bmatrix} a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and we can take $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as the eigenvector corresponding to $\lambda = 0$. Similarly, for $\lambda = 2$,

$$2\mathbf{I}_2 - A = \begin{bmatrix} 1 & 3 \\ 1 & 3 \end{bmatrix},$$

which has row-echelon form

$$\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}.$$

If

$$\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

then $a = -3b$. Thus

$$\begin{bmatrix} a \\ b \end{bmatrix} = a\begin{bmatrix} 1 \\ -3 \end{bmatrix}$$

and we can take $\begin{bmatrix} 1 \\ -3 \end{bmatrix}$ as the eigenvector corresponding to $\lambda = 2$.

7.2.5 ❶ We have

$$\begin{aligned}
\mathbf{det}(\lambda \mathbf{I}_3 - A) &= \mathbf{det}\begin{bmatrix} \lambda & -2 & 1 \\ -2 & \lambda - 3 & 2 \\ 1 & 2 & \lambda \end{bmatrix} \\
&= \lambda\,\mathbf{det}\begin{bmatrix} \lambda - 3 & 2 \\ 2 & \lambda \end{bmatrix} + 2\,\mathbf{det}\begin{bmatrix} -2 & 2 \\ 1 & \lambda \end{bmatrix} + \mathbf{det}\begin{bmatrix} -2 & \lambda - 3 \\ 1 & 2 \end{bmatrix} \\
&= \lambda(\lambda^2 - 3\lambda - 4) + 2(-2\lambda - 2) + (-\lambda - 1) \\
&= \lambda(\lambda - 4)(\lambda + 1) - 5(\lambda + 1) \\
&= (\lambda^2 - 4\lambda - 5)(\lambda + 1) \\
&= (\lambda + 1)^2(\lambda - 5)
\end{aligned}$$

❷ The eigenvalues are $-1, -1, 5$.

❸ If $\lambda = -1$,

$$(-\mathbf{I}_3 - A) = \begin{bmatrix} -1 & -2 & 1 \\ -2 & -4 & 2 \\ 1 & 2 & -1 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff a = -2b + c$$

$$\iff \begin{bmatrix} a \\ b \\ c \end{bmatrix} = b\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix} + c\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

We may take as eigenvectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, which are clearly linearly independent.

If $\lambda = 5$,

$$(5\mathbf{I}_3 - A) = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 2 & 2 \\ 1 & 2 & 5 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \iff a = -c, b = -2c,$$

$$\iff \begin{bmatrix} a \\ b \\ c \end{bmatrix} = c\begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}.$$

We may take as eigenvector $\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$.

7.2.6 The characteristic polynomial of A must be $\lambda^2 - 1$, which means that $\mathbf{tr}(A) = 0$ and $\mathbf{det}\, A = -1$. Hence A must be of the form $\begin{bmatrix} a & c \\ b & -a \end{bmatrix}$, with $-a^2 - bc = -1$, that is, $a^2 + bc = 1$.

7.2.7 We must shew that $\mathbf{det}(\lambda \mathbf{I}_n - A) = \mathbf{det}(\lambda \mathbf{I}_n - A)^T$. Now, recall that the determinant of a square matrix is the same as the determinant of its transpose. Hence

$$\mathbf{det}(\lambda \mathbf{I}_n - A) = \mathbf{det}((\lambda \mathbf{I}_n - A)^T) = \mathbf{det}(\lambda \mathbf{I}_n^T - A^T) = \mathbf{det}(\lambda \mathbf{I}_n - A^T),$$

as we needed to shew.

7.3.1 Put

$$D = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}, \quad , P = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}.$$

We find

$$P^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix}.$$

Since $A = PDP^{-1}$

$$A^{10} = PD^{10}P^{-1} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1024 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1023 \\ 0 & 1024 \end{bmatrix}.$$

7.3.2

1. A has characteristic polynomial $\mathbf{det} \begin{bmatrix} \lambda - 9 & 4 \\ -20 & \lambda + 9 \end{bmatrix} = (\lambda - 9)(\lambda + 9) + 80 = \lambda^2 - 1 = (\lambda - 1)(\lambda + 1)$.
2. $(\lambda - 1)(\lambda + 1) = 0 \implies \lambda \in \{-1, 1\}$.
3. For $\lambda = -1$ we have

$$\begin{bmatrix} 9 & -4 \\ 20 & -9 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = -1 \begin{bmatrix} a \\ b \end{bmatrix} \implies 10a = 4b \implies a = \frac{2b}{5},$$

so we can take $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$ as an eigenvector.

For $\lambda = 1$ we have

$$\begin{bmatrix} 9 & -4 \\ 20 & -9 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 1 \begin{bmatrix} a \\ b \end{bmatrix} \implies 8a = 4b \implies a = \frac{b}{2},$$

so we can take $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ as an eigenvector.

4. We can do this problem in at least three ways. The quickest is perhaps the following.

Recall that a 2×2 matrix has characteristic polynomial $\lambda^2 - (\mathbf{tr}\,(A))\lambda + \mathbf{det}\,A$. Since A has eigenvalues -1 and 1, A^{20} has eigenvalues $1^{20} = 1$ and $(-1)^{20} = 1$, i.e., the sole of A^{20} is 1 and so A^{20} has characteristic polynomial $(\lambda - 1)^2 = \lambda^2 - 2\lambda + 1$. This means that $-\mathbf{tr}\,(A^{20}) = -2$ and so $\mathbf{tr}\,(A^{20}) = 2$.

The direct way would be to argue that

$$\begin{aligned} A^{20} &= \begin{bmatrix} 2 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}^{20} \begin{bmatrix} 2 & 1 \\ 5 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 2 & 1 \\ 5 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 5 & 2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned}$$

and so $a + d = 2$. One may also use the fact that $\mathbf{tr}\,(XY) = \mathbf{tr}\,(YX)$ and hence

$$\mathbf{tr}\left(A^{20}\right) = \mathbf{tr}\left(PD^{20}P^{-1}\right) = \mathbf{tr}\left(PP^{-1}D^{20}\right) = \mathbf{tr}\left(D^{20}\right) = 2.$$

7.3.3 Put

$$D = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad X = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we know that $A = XDX^{-1}$ and so we need to find X^{-1}. But this is readily obtained by performing $R_1 - R_2 \to R_1$ and $R_2 - R_3 \to R_3$ in the augmented matrix

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right],$$

getting

$$X^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus

$$\begin{aligned} A &= XDX^{-1} \\ &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} -1 & 0 & 4 \\ 0 & -1 & 4 \\ 0 & 0 & 3 \end{bmatrix}. \end{aligned}$$

7.3.4 The determinant is 1, $A = A^{-1}$, and the characteristic polynomial is $(\lambda^2 - 1)^2$.

7.3.6 We find

$$\mathbf{det}(\lambda \mathbf{I}_2 - A) = \mathbf{det} \begin{bmatrix} \lambda + 7 & 6 \\ -12 & \lambda - 10 \end{bmatrix} = \lambda^2 - 3\lambda + 2 = (\lambda - 1)(\lambda - 2).$$

A short calculation shews that the eigenvalue $\lambda = 2$ has eigenvector $\begin{bmatrix} 2 \\ -3 \end{bmatrix}$ and that the eigenvalue $\lambda = 1$ has eigenvector $\begin{bmatrix} 3 \\ -4 \end{bmatrix}$. Thus we may form

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad P = \begin{bmatrix} 2 & -3 \\ 3 & -4 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} 4 & -3 \\ 3 & 1 \end{bmatrix}.$$

This gives

$$A = PDP^{-1} \implies A^n = PD^nP^{-1} = \begin{bmatrix} 2 & -3 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} 2^n & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & -3 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -8 \cdot 2^n + 9 & -6 \cdot 2^n + 6 \\ 12 \cdot 2^n - 12 & 9 \cdot 2^n - 8 \end{bmatrix}.$$

7.4.1 The eigenvalues of A are 0, 1, and -2. Those of A^2 are 0, 1, and 4. Hence, the characteristic polynomial of A^2 is $\lambda(\lambda - 1)(\lambda - 4)$.

8.2.1 $\sqrt{2a^2 - 2a + 1}$

8.2.2 $||\lambda \overrightarrow{v}|| = \frac{1}{2} \implies \sqrt{(\lambda)^2 + (-\lambda)^2} = \frac{1}{2} \implies 2\lambda^2 = \frac{1}{4} \implies \lambda = \pm \frac{1}{\sqrt{8}}$.

8.2.3 $\overrightarrow{0}$

8.2.4 $a = \pm 1$ or $a = -8$.

8.2.5 [A] $2(\overrightarrow{x} + \overrightarrow{y}) - \frac{1}{2}\overrightarrow{z}$, [B] $\overrightarrow{x} + \overrightarrow{y} - \frac{1}{2}\overrightarrow{z}$, [C] $-(\overrightarrow{x} + \overrightarrow{y} + \overrightarrow{z})$

8.2.6 [A]. $\overrightarrow{0}$, [B]. $\overrightarrow{0}$, [C]. $\overrightarrow{0}$, [D]. $\overrightarrow{0}$, [E]. $2\overrightarrow{c}\,(= 2\overrightarrow{d})$

8.2.7 [F]. $\overrightarrow{0}$, [G]. $\overrightarrow{b}$, [H]. $2\overrightarrow{0}$, [I]. $\overrightarrow{0}$.

8.2.8 Let the skew quadrilateral be $ABCD$ and let P, Q, R, S be the midpoints of $[A, B]$, $[B, C]$, $[C, D]$, $[D, A]$, respectively. Put $\overrightarrow{x} = \overrightarrow{OX}$, where $X \in \{A, B, C, D, P, Q, R, S\}$. Using the Section Formula 8.4 we have

$$\overrightarrow{p} = \frac{\overrightarrow{a} + \overrightarrow{b}}{2}, \quad \overrightarrow{q} = \frac{\overrightarrow{b} + \overrightarrow{c}}{2}, \quad \overrightarrow{r} = \frac{\overrightarrow{c} + \overrightarrow{d}}{2}, \quad \overrightarrow{s} = \frac{\overrightarrow{d} + \overrightarrow{a}}{2}.$$

This gives

$$\overrightarrow{p} - \overrightarrow{q} = \frac{\overrightarrow{a} - \overrightarrow{c}}{2}, \quad \overrightarrow{s} - \overrightarrow{r} = \frac{\overrightarrow{a} - \overrightarrow{c}}{2}.$$

This means that $\overrightarrow{QP} = \overrightarrow{RS}$ and so $PQRS$ is a parallelogram since one pair of sides are equal and parallel.

8.2.9 We have $2\overrightarrow{BC} = \overrightarrow{BE} + \overrightarrow{EC}$. By Chasles' Rule $\overrightarrow{AC} = \overrightarrow{AE} + \overrightarrow{EC}$, and $\overrightarrow{BD} = \overrightarrow{BE} + \overrightarrow{ED}$. We deduce that

$$\overrightarrow{AC} + \overrightarrow{BD} = \overrightarrow{AE} + \overrightarrow{EC} + \overrightarrow{BE} + \overrightarrow{ED} = \overrightarrow{AD} + \overrightarrow{BC}.$$

But since $ABCD$ is a parallelogram, $\overrightarrow{AD} = \overrightarrow{BC}$. Hence

$$\overrightarrow{AC} + \overrightarrow{BD} = \overrightarrow{AD} + \overrightarrow{BC} = 2\overrightarrow{BC}.$$

8.2.10 We have $\overrightarrow{IA} = -3\overrightarrow{IB} \iff \overrightarrow{IA} = -3(\overrightarrow{IA} + \overrightarrow{AB}) = -3\overrightarrow{IA} - 3\overrightarrow{AB}$. Thus we deduce

$$\begin{aligned} \overrightarrow{IA} + 3\overrightarrow{IA} = -3\overrightarrow{AB} &\iff 4\overrightarrow{IA} = -3\overrightarrow{AB} \\ &\iff 4\overrightarrow{AI} = 3\overrightarrow{AB} \\ &\iff \overrightarrow{AI} = \tfrac{3}{4}\overrightarrow{AB}. \end{aligned}$$

Similarly

$$\begin{aligned} \overrightarrow{JA} = -\tfrac{1}{3}\overrightarrow{JB} &\iff 3\overrightarrow{JA} = -\overrightarrow{JB} \\ &\iff 3\overrightarrow{JA} = -\overrightarrow{JA} - \overrightarrow{AB} \\ &\iff 4\overrightarrow{JA} = -\overrightarrow{AB} \\ &\iff \overrightarrow{AJ} = \tfrac{1}{4}\overrightarrow{AB} \end{aligned}$$

.

Thus we take I such that $\overrightarrow{AI} = \frac{3}{4}\overrightarrow{AB}$ and J such that $\overrightarrow{AJ} = \frac{1}{4}\overrightarrow{AB}$.

Now

$$\begin{aligned} \overrightarrow{MA} + 3\overrightarrow{MB} &= \overrightarrow{MI} + \overrightarrow{IA} + 3\overrightarrow{IB} \\ &= 4\overrightarrow{MI} + \overrightarrow{IA} + 3\overrightarrow{IB} \\ &= 4\overrightarrow{MI}, \end{aligned}$$

and

$$\begin{aligned} 3\overrightarrow{MA} + \overrightarrow{MB} &= 3\overrightarrow{MJ} + 3\overrightarrow{JA} + \overrightarrow{MJ} + \overrightarrow{JB} \\ &= 4\overrightarrow{MJ} + 3\overrightarrow{JA} + \overrightarrow{JB} \\ &= 4\overrightarrow{MJ}. \end{aligned}$$

8.2.11 Let $\overrightarrow{G}$, $\overrightarrow{O}$ and $\overrightarrow{P}$ denote vectors from an arbitrary origin to the gallows, oak, and pine, respectively. The conditions of the problem define $\overrightarrow{X}$ and $\overrightarrow{Y}$, thought of similarly as vectors from the origin, by $\overrightarrow{X} = \overrightarrow{O} + R(\overrightarrow{O} - \overrightarrow{G})$, $\overrightarrow{Y} = \overrightarrow{P} - R(\overrightarrow{P} - \overrightarrow{G})$, where R is the $90°$ rotation to the right, a linear transformation on vectors in the plane; the fact that $-R$ is $90°$ leftward rotation has been used in writing Y. Anyway, then

$$\frac{\overrightarrow{X} + \overrightarrow{Y}}{2} = \frac{\overrightarrow{O} + \overrightarrow{P}}{2} + \frac{R(\overrightarrow{O} - \overrightarrow{P})}{2}$$

is independent of the position of the gallows. This gives a simple algorithm for treasure-finding: take $\overrightarrow{P}$ as the (hitherto) arbitrary origin, then the treasure is at $\dfrac{\overrightarrow{O} + R(\overrightarrow{O})}{2}$.

8.3.1 $a = \frac{1}{2}$

8.3.3

$$\overrightarrow{p} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} = 2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} + 3 \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2\overrightarrow{r} + 3\overrightarrow{s}.$$

8.3.4 Since $a_1 = \overrightarrow{a} \bullet \overrightarrow{i}$, $a_2 = \overrightarrow{a} \bullet \overrightarrow{j}$, we may write

$$\overrightarrow{a} = (\overrightarrow{a} \bullet \overrightarrow{i})\overrightarrow{i} + (\overrightarrow{a} \bullet \overrightarrow{j})\overrightarrow{j}$$

from where the assertion follows.

8.3.5

$$\begin{aligned} \alpha\overrightarrow{a} + \beta\overrightarrow{b} = \overrightarrow{0} \quad &\implies \quad \overrightarrow{a} \bullet (\alpha\overrightarrow{a} + \beta\overrightarrow{b}) = \overrightarrow{a} \bullet \overrightarrow{0} \\ &\implies \quad \alpha(\overrightarrow{a} \bullet \overrightarrow{a}) = 0 \\ &\implies \quad \alpha \left|\left|\overrightarrow{a}\right|\right|^2 = 0. \end{aligned}$$

Since $\overrightarrow{a} \neq \overrightarrow{0}$, we must have $\left|\left|\overrightarrow{a}\right|\right| \neq 0$ and thus $\alpha = 0$. But if $\alpha = 0$ then

$$\begin{aligned} \alpha\overrightarrow{a} + \beta\overrightarrow{b} = \overrightarrow{0} \quad &\implies \quad \beta\overrightarrow{b} = \overrightarrow{0} \\ &\implies \quad \beta = 0, \end{aligned}$$

since $\overrightarrow{b} \neq \overrightarrow{0}$.

8.3.6 We must shew that

$$(2\overrightarrow{x} + 3\overrightarrow{y}) \bullet (2\overrightarrow{x} - 3\overrightarrow{y}) = 0.$$

But

$$(2\overrightarrow{x} + 3\overrightarrow{y}) \bullet (2\overrightarrow{x} - 3\overrightarrow{y}) = 4||\overrightarrow{x}||^2 - 9||\overrightarrow{y}||^2 = 4(\frac{9}{4}||\overrightarrow{y}||^2) - 9||\overrightarrow{y}||^2 = 0.$$

8.3.7 We have $\forall \overrightarrow{v} \in \mathbb{R}^2, \overrightarrow{v} \bullet (\overrightarrow{a} - \overrightarrow{b}) = 0$. In particular, choosing $\overrightarrow{v} = \overrightarrow{a} - \overrightarrow{b}$, we gather

$$(\overrightarrow{a} - \overrightarrow{b}) \bullet (\overrightarrow{a} - \overrightarrow{b}) = ||\overrightarrow{a} - \overrightarrow{b}||^2 = 0.$$

But the norm of a vector is 0 if and only if the vector is the $\overrightarrow{0}$ vector. Therefore $\overrightarrow{a} - \overrightarrow{b} = \overrightarrow{0}$, i.e., $\overrightarrow{a} = \overrightarrow{b}$.

8.3.8 We have

$$\begin{aligned} \left|\left|\overrightarrow{a} \pm \overrightarrow{b}\right|\right|^2 &= (\overrightarrow{a} \pm \overrightarrow{b}) \bullet (\overrightarrow{a} \pm \overrightarrow{b}) \\ &= \overrightarrow{a} \bullet \overrightarrow{a} \pm 2\overrightarrow{a} \bullet \overrightarrow{b} + \overrightarrow{b} \bullet \overrightarrow{b} \\ &= \left|\left|\overrightarrow{a}\right|\right|^2 \pm 2\overrightarrow{a} \bullet \overrightarrow{b} + \left|\left|\overrightarrow{b}\right|\right|^2, \end{aligned}$$

whence the result follows.

8.3.9 We have

$$\begin{aligned} ||\overrightarrow{u} + \overrightarrow{v}||^2 - ||\overrightarrow{u} - \overrightarrow{v}||^2 &= (\overrightarrow{u} + \overrightarrow{v}) \bullet (\overrightarrow{u} + \overrightarrow{v}) - (\overrightarrow{u} - \overrightarrow{v}) \bullet (\overrightarrow{u} - \overrightarrow{v}) \\ &= \overrightarrow{u} \bullet \overrightarrow{u} + 2\overrightarrow{u} \bullet \overrightarrow{v} + \overrightarrow{v} \bullet \overrightarrow{v} - (\overrightarrow{u} \bullet \overrightarrow{u} - 2\overrightarrow{u} \bullet \overrightarrow{v} + \overrightarrow{v} \bullet \overrightarrow{v}) \\ &= 4\overrightarrow{u} \bullet \overrightarrow{v}, \end{aligned}$$

giving the result.

8.3.10 By definition

$$\begin{aligned}\mathbf{proj}_{\vec{a}}^{\mathbf{proj}_{\vec{x}}^{\vec{a}}} &= \frac{\mathbf{proj}_{\vec{x}}^{\vec{a}} \bullet \vec{a}}{||\vec{a}||^2}\vec{a} \\ &= \frac{\frac{\vec{a}\bullet\vec{x}}{||\vec{x}||^2}\vec{x}\bullet\vec{a}}{||\vec{a}||^2}\vec{a} \\ &= \frac{(\vec{a}\bullet\vec{x})^2}{||\mathbf{x}||^2||\vec{a}||^2}\vec{a},\end{aligned}$$

Since $0 \leq \frac{(\vec{a}\bullet\vec{x})^2}{||\vec{x}||^2||\vec{a}||^2} \leq 1$ by the CBS Inequality, the result follows.

8.3.11 Clearly, if $\vec{a} = \vec{\mathbf{0}}$ and $\lambda \neq 0$ then there are no solutions. If both $\vec{a} = \vec{\mathbf{0}}$ and $\lambda = 0$, then the solution set is the whole space $\mathbb{R}^2$. So assume that $\vec{a} \neq \vec{\mathbf{0}}$. By Theorem 365, we may write $\vec{x} = \vec{u} + \vec{v}$ with $\mathbf{proj}_{\vec{\mathbf{a}}}^{\vec{\mathbf{x}}} = \vec{u} || \vec{a}$ and $\vec{v} \perp \vec{a}$. Thus there are infinitely many solutions, each of the form

$$\vec{x} = \vec{u} + \vec{v} = \frac{\vec{x}\bullet\vec{a}}{||\mathbf{a}||^2}\vec{a} + \vec{v} = \frac{\lambda}{||\mathbf{a}||^2}\vec{a} + \vec{v},$$

where $\vec{v} \in \vec{a}^{\perp}$.

8.4.1 Since $\vec{a} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ is normal to $2x - y = 1$ and $\vec{b} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$ is normal to $x - 3y = 1$, the desired angle can be obtained by finding the angle between the normal vectors:

$$\widehat{(\vec{a},\vec{b})} = \mathbf{arccos}\,\frac{\vec{a}\bullet\vec{b}}{||\vec{a}||||\vec{b}||} = \mathbf{arccos}\,\frac{5}{\sqrt{5}\cdot\sqrt{10}} = \mathbf{arccos}\,\frac{1}{\sqrt{2}} = \frac{\pi}{4}.$$

8.4.2 $2(x-1) + (y+1) = 0$ or $2x + y = 1$.

8.4.3 By Chasles' Rule $\overrightarrow{AA'} = \overrightarrow{AG} + \overrightarrow{GA'}$, $\overrightarrow{BB'} = \overrightarrow{BG} + \overrightarrow{GB'}$, and $\overrightarrow{CC'} = \overrightarrow{CG} + \overrightarrow{GC'}$. Thus

$$\begin{aligned}\vec{0} &= \overrightarrow{AA'} + \overrightarrow{BB'} + \overrightarrow{CC'} \\ &= \overrightarrow{AG} + \overrightarrow{GA'} + \overrightarrow{BG} + \overrightarrow{GB'} + \overrightarrow{CG} + \overrightarrow{GC'} \\ &= -(\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC}) + (\overrightarrow{GA'} + \overrightarrow{GB'} + \overrightarrow{GC'}) \\ &= \overrightarrow{GA'} + \overrightarrow{GB'} + \overrightarrow{GC'},\end{aligned}$$

whence the result.

8.4.4 We have:

❶ The points F, A, D are collinear, and so $\overrightarrow{FA}$ is parallel to $\overrightarrow{FD}$, meaning that there is $k \in \mathbb{R} \setminus \{0\}$ such that $\overrightarrow{FA} = k\overrightarrow{FD}$. Since the lines (AB) and (DC) are parallel, we obtain through Thales' Theorem that $\overrightarrow{FI} = k\overrightarrow{FJ}$ and $\overrightarrow{FB} = k\overrightarrow{FC}$. This gives

$$\overrightarrow{FA} - \overrightarrow{FI} = k(\overrightarrow{FD} - \overrightarrow{FJ}) \implies \overrightarrow{IA} = k\overrightarrow{JD}.$$

Similarly

$$\overrightarrow{FB} - \overrightarrow{FI} = k(\overrightarrow{FC} - \overrightarrow{FJ}) \implies \overrightarrow{IB} = k\overrightarrow{JC}.$$

Since I is the midpoint of $[A, B]$, $\overrightarrow{IA} + \overrightarrow{IB} = \vec{0}$, and thus $k(\overrightarrow{JC} + \overrightarrow{JD}) = \vec{0}$. Since $k \neq 0$, we have $\overrightarrow{JC} + \overrightarrow{JD} = \vec{0}$, meaning that J is the midpoint of $[C, D]$. Therefore the midpoints of $[A, B]$ and $[C, D]$ are aligned with F.

❷ Let J' be the intersection of the lines (EI) and (DC). Let us prove that $J' = J$.

Since the points E, A, C are collinear, there is $l \neq 0$ such that $\overrightarrow{EA} = l\overrightarrow{EC}$. Since the lines (ab) and (DC) are parallel, we obtain via Thales' Theorem that $\overrightarrow{EI} = l\overrightarrow{EJ'}$ and $\overrightarrow{EB} = l\overrightarrow{ED}$. These equalities give

$$\overrightarrow{EA} - \overrightarrow{EI} = l(\overrightarrow{EC} - \overrightarrow{EJ'}) \implies \overrightarrow{IA} = l\overrightarrow{J'C},$$

$$\overrightarrow{EB} - \overrightarrow{EI} = l(\overrightarrow{ED} - \overrightarrow{EJ'}) \implies \overrightarrow{IB} = l\overrightarrow{J'D}.$$

Since I is the midpoint of $[A, B]$, $\overrightarrow{IA}+\overrightarrow{IB} = \overrightarrow{0}$, and thus $l(\overrightarrow{J'C}+\overrightarrow{J'D}) = \overrightarrow{0}$. Since $l \neq 0$, we deduce $\overrightarrow{J'C}+\overrightarrow{J'D} = \overrightarrow{0}$, that is, J' is the midpoint of $[C, D]$, and so $J' = J$.

8.4.5 We have:

❶ By Chasles' Rule

$$\overrightarrow{AE} = \tfrac{1}{4}\overrightarrow{AC} \iff \overrightarrow{AB} + \overrightarrow{BE} = \tfrac{1}{4}\overrightarrow{AC} ,$$

and

$$\overrightarrow{AF} = \tfrac{3}{4}\overrightarrow{AC} \iff \overrightarrow{AD} + \overrightarrow{DF} = \tfrac{3}{4}\overrightarrow{AC} .$$

Adding, and observing that since $ABCD$ is a parallelogram, $\overrightarrow{AB} = \overrightarrow{CD}$,

$$\begin{aligned}\overrightarrow{AB} + \overrightarrow{BE} + \overrightarrow{AD} + \overrightarrow{DF} = \overrightarrow{AC} &\iff \overrightarrow{BE} + \overrightarrow{DF} = \overrightarrow{AC} - \overrightarrow{AB} - \overrightarrow{AD} \\ &\iff \overrightarrow{BE} + \overrightarrow{DF} = \overrightarrow{AD} + \overrightarrow{DC} - \overrightarrow{AB} - \overrightarrow{AD} . \\ &\iff \overrightarrow{BE} = -\overrightarrow{DF}.\end{aligned}$$

The last equality shews that the lines (BE) and (DF) are parallel.

❷ Observe that $\overrightarrow{BJ} = \tfrac{1}{2}\overrightarrow{BC} = \tfrac{1}{2}\overrightarrow{AD} = \overrightarrow{AI} = -\overrightarrow{IA}$. Hence

$$\overrightarrow{IJ} = \overrightarrow{IA} + \overrightarrow{AB} + \overrightarrow{BJ} = \overrightarrow{AB},$$

proving that the lines (AB) and (IJ) are parallel.

Observe that

$$\overrightarrow{IE} = \overrightarrow{IA} + \overrightarrow{AE} = \frac{1}{2}\overrightarrow{DA} + \frac{1}{4}\overrightarrow{AC} = \frac{1}{2}\overrightarrow{CB} + \overrightarrow{FC} = \overrightarrow{CJ} + \overrightarrow{FC} = \overrightarrow{FC} + \overrightarrow{CJ} = \overrightarrow{FJ},$$

whence $IEJF$ is a parallelogram.

8.4.6 Since $\overrightarrow{IE} = \tfrac{1}{3}\overrightarrow{ID}$ and $[I, D]$ is a median of $\triangle ABD$, E is the centre of gravity of $\triangle ABD$. Let M be the midpoint of $[B, D]$, and observe that M is the centre of the parallelogram, and so $2\overrightarrow{AM} = \overrightarrow{AB} + \overrightarrow{AD}$. Thus

$$\overrightarrow{AE} = \frac{2}{3}\overrightarrow{AM} = \frac{1}{3}(2\overrightarrow{AM}) = \frac{1}{3}(\overrightarrow{AB} + \overrightarrow{AD}).$$

To shew that A, C, E are collinear it is enough to notice that $\overrightarrow{AE} = \tfrac{1}{3}\overrightarrow{AC}$.

8.4.7 Suppose A, B, C are collinear and that $\dfrac{\|[A, B]\|}{\|[B, C]\|} = \dfrac{\lambda}{\mu}$. Then by the Section Formula 8.4,

$$\overrightarrow{b} = \frac{\lambda\overrightarrow{c} + \mu\overrightarrow{a}}{\lambda + \mu},$$

whence $\mu\overrightarrow{a} - (\lambda + \mu)\overrightarrow{b} + \lambda\overrightarrow{c} = \overrightarrow{0}$ and clearly $\mu - (\lambda + \mu) + \lambda = 0$. Thus we may take $\alpha = \mu$, $\beta = \lambda + \mu$, and $\gamma = \lambda$. Conversely, suppose that

$$\alpha\overrightarrow{a} + \beta\overrightarrow{b} + \gamma\overrightarrow{c} = \overrightarrow{0}, \quad \alpha + \beta + \gamma = 0$$

for some real numbers α, β, γ, not all zero. Assume without loss of generality that $\gamma \neq 0$. Otherwise we simply change the roles of γ, and α and β Then $\gamma = -(\alpha + \beta) \neq 0$. Hence

$$\alpha\overrightarrow{a} + \beta\overrightarrow{b} = (\alpha + \beta)\overrightarrow{c} \implies \overrightarrow{c} = \frac{\alpha\overrightarrow{a} + \beta\overrightarrow{b}}{\alpha + \beta},$$

and thus $[O, C]$ divides $[A, B]$ into the ratio $\dfrac{\beta}{\alpha}$, and therefore, A, B, C are collinear.

8.4.8 Put $\overrightarrow{OX} = \overrightarrow{x}$ for $X \in \{A, A', B, B', C, C', L, M, N, V\}$. Using problem 8.4.7 we deduce

$$\overrightarrow{v} + \alpha\overrightarrow{a} + \alpha'\overrightarrow{a'} = \overrightarrow{0}, \quad 1 + \alpha + \alpha' = 0, \tag{A.1}$$

$$\overrightarrow{v} + \beta\overrightarrow{a} + \beta'\overrightarrow{a'} = \overrightarrow{0}, \quad 1 + \beta + \beta' = 0, \tag{A.2}$$

$$\overrightarrow{v} + \gamma\overrightarrow{a} + \gamma'\overrightarrow{a'} = \overrightarrow{0}, \quad 1 + \gamma + \gamma' = 0. \tag{A.3}$$

From A.2, A.3, and the Section Formula 8.4 we find

$$\frac{\beta\overrightarrow{b} - \gamma\overrightarrow{c}}{\beta - \gamma} = \frac{\beta'\overrightarrow{b'} - \gamma'\overrightarrow{c'}}{\beta' - \gamma'} = \overrightarrow{l},$$

whence $(\beta - \gamma)\overrightarrow{l} = \beta\overrightarrow{b} - \gamma\overrightarrow{c}$. In a similar fashion, we deduce

$$(\gamma - \alpha)\overrightarrow{m} = \gamma\overrightarrow{c} - \alpha\overrightarrow{a},$$

$$(\alpha - \beta)\overrightarrow{n} = \alpha\overrightarrow{a} - \beta\overrightarrow{b}.$$

This gives

$$(\beta - \gamma)\overrightarrow{l} + (\gamma - \alpha)\overrightarrow{m} + (\alpha - \beta)\overrightarrow{n} = \overrightarrow{0},$$
$$(\beta - \gamma) + (\gamma - \alpha) + (\alpha - \beta) = 0,$$

and appealing to problem 8.4.7 once again, we deduce that L, M, N are collinear.

8.5.1 [A] $\overrightarrow{AS}$, [B] $\overrightarrow{AB}$.

8.5.2 Put

$$\overrightarrow{a} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = (\overrightarrow{i} + \overrightarrow{j} + \overrightarrow{k}) \times (\overrightarrow{i} + \overrightarrow{j}) = \overrightarrow{j} - \overrightarrow{i} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Then either

$$\frac{3\overrightarrow{a}}{\|\overrightarrow{a}\|} = \frac{3\overrightarrow{a}}{\sqrt{2}} = \begin{bmatrix} -\frac{3}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} \\ 0 \end{bmatrix},$$

or

$$-\frac{3\overrightarrow{a}}{\|\overrightarrow{a}\|} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ -\frac{3}{\sqrt{2}} \\ 0 \end{bmatrix}$$

will satisfy the requirements.

8.5.3 The desired area is

$$\left\|\overrightarrow{PQ} \times \overrightarrow{PR}\right\| = \left\|\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}\right\| = \left\|\begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}\right\| = \sqrt{3}.$$

8.5.4 It is not associative, since $\overrightarrow{i} \times (\overrightarrow{i} \times \overrightarrow{j}) = \overrightarrow{i} \times \overrightarrow{k} = -\overrightarrow{j}$ but $(\overrightarrow{i} \times \overrightarrow{i}) \times \overrightarrow{j} = \overrightarrow{0} \times \overrightarrow{j} = \mathbf{0}$.

8.5.5 We have $\overrightarrow{x} \times \overrightarrow{x} = -\overrightarrow{x} \times \overrightarrow{x}$ by letting $\overrightarrow{y} = \overrightarrow{x}$ in 8.15. Thus $2\overrightarrow{x} \times \overrightarrow{x} = \overrightarrow{0}$ and hence $\overrightarrow{x} \times \overrightarrow{x} = \overrightarrow{0}$.

8.5.6 $2\overrightarrow{a} \times \overrightarrow{b}$

8.5.7

$$\vec{a}\times(\vec{x}\times\vec{b}) = \vec{b}\times(\vec{x}\times\vec{a}) \iff (\vec{a}\bullet\vec{b})\vec{x} - (\vec{a}\bullet\vec{x})\vec{b} = (\vec{b}\bullet\vec{a})\vec{x} - (\vec{b}\bullet\vec{x})\vec{a} \iff \vec{a}\bullet\vec{x} = \vec{b}\bullet\vec{x} = 0.$$

The answer is thus $\{\vec{x} : \vec{x} \in \mathbb{R}\vec{a}\times\vec{b}\}$.

8.5.8

$$\vec{x} = \frac{(\vec{a}\bullet\vec{b})\vec{a} + 6\vec{b} + 2\vec{a}\times\vec{c}}{12 + 2\left|\left|\vec{a}\right|\right|^2}$$

$$\vec{y} = \frac{(\vec{a}\bullet\vec{c})\vec{a} + 6\vec{c} + 3\vec{a}\times\vec{b}}{18 + 3\left|\left|\vec{a}\right|\right|^2}$$

8.5.9 Assume contrariwise that $\vec{a}$, $\vec{b}$, $\vec{c}$ are three unit vectors in $\mathbb{R}^3$ such that the angle between any two of them is $> \frac{2\pi}{3}$. Then $\vec{a}\bullet\vec{b} < -\frac{1}{2}$, $\vec{b}\bullet\vec{c} < -\frac{1}{2}$, and $\vec{c}\bullet\vec{a} < -\frac{1}{2}$. Thus

$$\begin{aligned}\left|\left|\vec{a} + \vec{b} + \vec{c}\right|\right|^2 &= \left|\left|\vec{a}\right|\right|^2 + \left|\left|\vec{b}\right|\right|^2 + \left|\left|\vec{c}\right|\right|^2 \\ &\quad +2\vec{a}\bullet\vec{b} + 2\vec{b}\bullet\vec{c} + 2\vec{c}\bullet\vec{a} \\ &< 1 + 1 + 1 - 1 - 1 - 1 \\ &= 0,\end{aligned}$$

which is impossible, since a norm of vectors is always ≥ 0.

8.5.10 Take $(\vec{u}, \vec{v}) \in X^2$ and $\alpha \in \mathbb{R}$. Then

$$\vec{a}\times(\vec{u} + \alpha\vec{v}) = \vec{a}\times\vec{u} + \alpha\vec{a}\times\vec{v} = \vec{0} + \alpha\vec{0} = \vec{0},$$

proving that X is a vector subspace of $\mathbb{R}^n$.

8.5.11 Since $\vec{a}, \vec{b}$ are linearly independent, none of them is $\vec{0}$. Assume that there are $(\alpha, \beta, \gamma) \in \mathbb{R}^3$ such that

$$\alpha\vec{a} + \beta\vec{b} + \gamma\vec{a}\times\vec{b} = \vec{0}. \tag{A.4}$$

Since $\vec{a}\bullet(\vec{a}\times\vec{b}) = 0$, taking the dot product of A.4 with $\vec{a}$ yields $\alpha\left|\left|\vec{a}\right|\right|^2 = 0$, which means that $\alpha = 0$, since $\left|\left|\vec{a}\right|\right| \neq 0$. Similarly, we take the dot product with $\vec{b}$ and $\vec{a}\times\vec{b}$ obtaining respectively, $\beta = 0$ and $\gamma = 0$. This establishes linear independence.

8.5.12 Since $\vec{a} \perp \vec{a}\times\vec{\mathbf{x}} = \vec{b}$, there are no solutions if $\vec{\mathbf{a}}\bullet\vec{\mathbf{b}} \neq 0$. Neither are there solutions if $\vec{a} = \vec{\mathbf{0}}$ and $\vec{b} \neq \vec{\mathbf{0}}$. If both $\vec{a} = \vec{b} = \vec{\mathbf{0}}$, then the solution set is the whole of $\mathbb{R}^3$. Assume thus that $\vec{\mathbf{a}}\bullet\vec{\mathbf{b}} = 0$ and that $\vec{a}$ and $\vec{b}$ are linearly independent. Then $\vec{a}, \vec{b}, \vec{a}\times\vec{b}$ are linearly independent, and so they constitute a basis for $\mathbb{R}^3$. Any $\vec{x} \in \mathbb{R}^3$ can be written in the form

$$\vec{x} = \alpha\vec{a} + \beta\vec{b} + \gamma\vec{a}\times\vec{b}.$$

We then have

$$\begin{aligned}\vec{b} &= \vec{a}\times\vec{x} \\ &= \beta\vec{a}\times\vec{b} + \gamma\vec{a}\times(\vec{a}\times\vec{b}) \\ &= \beta\vec{a}\times\vec{b} + \gamma((\vec{a}\bullet\vec{b})\vec{a} - (\vec{a}\bullet\vec{a})\vec{b}). \\ &= \beta\vec{a}\times\vec{b} - \gamma(\vec{a}\bullet\vec{a}\vec{b}) \\ &= \beta\vec{a}\times\vec{b} - \gamma\left|\left|\vec{a}\right|\right|^2\vec{b},\end{aligned}$$

from where

$$\beta\vec{a}\times\vec{\mathbf{b}} + (-\gamma\left|\left|\vec{a}\right|\right|^2 - 1)\vec{b} = \vec{0},$$

which means that $\beta = 0$ and $\gamma = -\frac{1}{\left|\left|\vec{a}\right|\right|^2}$, since $\vec{a}, \vec{b}, \vec{a}\times\vec{b}$ are linearly independent. Thus

$$\vec{x} = \alpha\vec{a} - \frac{1}{\left|\left|\vec{a}\right|\right|^2}\vec{a}\times\vec{b}$$

in this last case.

8.5.13 Let $\vec{x}, \vec{y}, \vec{x}', \vec{y}'$ be vectors in $\mathbb{R}^3$ and let $\alpha \in \mathbb{R}$ be a scalar. Then

$$\begin{aligned} L((\vec{x},\vec{y}) + \alpha(\vec{x}',\vec{y}')) &= L(\vec{x} + \alpha\vec{x}', \vec{y} + \alpha\vec{y}') \\ &= (\vec{x} + \alpha\vec{x}')\times\vec{k} + \vec{h}\times(\vec{y} + \alpha\vec{y}') \\ &= \vec{x}\times\vec{k} + \alpha\vec{x}'\times\vec{k} + \vec{h}\times\vec{y} + \vec{h}\times\alpha\vec{y}' \\ &= L(\vec{x},\vec{y}) + \alpha L(\vec{x}',\vec{y}') \end{aligned}$$

8.6.1 The vectors

$$\begin{bmatrix} a-(-a) \\ 0-1 \\ a-0 \end{bmatrix} = \begin{bmatrix} 2a \\ -1 \\ a \end{bmatrix}$$

and

$$\begin{bmatrix} 0-(-a) \\ 1-1 \\ 2a-0 \end{bmatrix} = \begin{bmatrix} a \\ 0 \\ 2a \end{bmatrix}$$

are coplanar. A vector normal to the plane is

$$\begin{bmatrix} 2a \\ -1 \\ a \end{bmatrix} \times \begin{bmatrix} a \\ 0 \\ 2a \end{bmatrix} = \begin{bmatrix} -2a \\ -3a^2 \\ a \end{bmatrix}.$$

The equation of the plane is thus given by

$$\begin{bmatrix} -2a \\ -3a^2 \\ a \end{bmatrix} \bullet \begin{bmatrix} x-a \\ y-0 \\ z-a \end{bmatrix} = 0,$$

that is,

$$2ax + 3a^2y - az = a^2.$$

8.6.2 The vectorial form of the equation of the line is

$$\vec{r} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

Since the line follows the direction of $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$, this means that $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ is normal to the plane, and thus the equation of the desired plane is

$$(x-1) - 2(y-1) - (z-1) = 0.$$

8.6.3 Observe that $(0, 0, 0)$ (as $0 = 2(0) = 3(0)$) is on the line, and hence on the plane. Thus the vector

$$\begin{bmatrix} 1-0 \\ -1-0 \\ -1-0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$$

lies on the plane. Now, if $x = 2y = 3z = t$, then $x = t, y = t/2, z = t/3$. Hence, the vectorial form of the equation of the line is

$$\overrightarrow{r} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix} = t \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix}.$$

This means that $\begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix}$ also lies on the plane, and thus

$$\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1/2 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 1/6 \\ -4/3 \\ 3/2 \end{bmatrix}$$

is normal to the plane. The desired equation is thus

$$\frac{1}{6}x - \frac{4}{3}y + \frac{3}{2}z = 0.$$

8.6.4 Put $ax = by = cz = t$, so $x = t/a; y = t/b; z = t/c$. The parametric equation of the line is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix}, \quad t \in \mathbb{R}.$$

Thus the vector $\begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix}$ is perpendicular to the plane. Therefore, the equation of the plane is

$$\begin{bmatrix} 1/a \\ 1/b \\ 1/c \end{bmatrix} \bullet \begin{bmatrix} x-1 \\ y-1 \\ z-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

or

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = \frac{1}{a} + \frac{1}{b} + \frac{1}{c}.$$

We may also write this as

$$bcx + cay + abz = ab + bc + ca.$$

8.6.5 A vector normal to the plane is $\begin{bmatrix} a \\ a^2 \\ a^2 \end{bmatrix}$. The line sought has the same direction as this vector, thus the equation of the line is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} a \\ a^2 \\ a^2 \end{bmatrix}, \quad t \in \mathbb{R}.$$

8.6.6 We have

$$x - z - y = 1 \implies -1 - y = 1 \implies y = -2.$$

Hence if $z = t$,

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} t-1 \\ -2 \\ t \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

8.6.7 The vector

$$\begin{bmatrix} 2-1 \\ 1-0 \\ 1-(-1) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

lies on the plane. The vector

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}$$

is normal to the plane. Hence the equation of the plane is

$$\begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \bullet \begin{bmatrix} x-1 \\ y \\ z+1 \end{bmatrix} = 0 \implies x + y - z = 2.$$

8.6.8 We have $\vec{c} \times \vec{a} = -\vec{i} + 2\vec{j}$ and $\vec{a} \times \vec{b} = 2\vec{k} - 3\vec{i}$. By Theorem 398, we have

$$\vec{b} \times \vec{c} = -\vec{a} \times \vec{b} - \vec{c} \times \vec{a} = -2\vec{k} + 3\vec{i} + \vec{i} - 2\vec{j} = 4\vec{i} - 2\vec{j} - 2\vec{k}.$$

8.6.9 $4x + 6y = 1$

8.6.10 There are 7 vertices ($V_0 = (0,0,0)$, $V_1 = (11,0,0)$, $V_2 = (0,9,0)$, $V_3 = (0,0,8)$, $V_4 = (0,3,8)$, $V_5 = (9,0,2)$, $V_6 = (4,7,0)$) and 11 edges (V_0V_1, V_0V_2, V_0V_3, V_1V_5, V_1V_6, V_2V_4, V_3V_4, V_3V_5, V_4V_5, and V_4V_6).

8.7.2 Expand $\left|\left|\sum_{i=1}^{n} \vec{a_i}\right|\right|^2 = 0$.

8.7.3 Observe that $\sum_{k=1}^{n} 1 = n$. Then we have

$$n^2 = \left(\sum_{k=1}^{n} 1\right)^2 = \left(\sum_{k=1}^{n} (a_k)\left(\frac{1}{a_k}\right)\right)^2 \leq \left(\sum_{k=1}^{n} a_k^2\right)\left(\sum_{k=1}^{n} \frac{1}{a_k^2}\right),$$

giving the result.

8.7.4 Take $(\vec{u}, \vec{v}) \in X^2$ and $\alpha \in \mathbb{R}$. Then

$$\vec{a} \bullet (\vec{u} + \alpha\vec{v}) = \vec{a} \bullet \vec{u} + \alpha \vec{a} \bullet \vec{v} = 0 + 0 = 0,$$

proving that X is a vector subspace of $\mathbb{R}^n$.

8.7.5 Assume that

$$\lambda_1 \vec{a_1} + \cdots + \lambda_k \vec{a_k} = \vec{0}.$$

Taking the dot product with $\vec{a_j}$ and using the fact that $\vec{a_i} \bullet \vec{a_j} = 0$ for $i \neq j$ we obtain

$$0 = \vec{0} \bullet \vec{a_j} = \lambda_j \vec{a_j} \bullet \vec{a_j} = \lambda_j ||a_j||^2.$$

Since $\vec{a_j} \neq \vec{0} \implies ||a_j||^2 \neq 0$, we must have $\lambda_j = 0$. Thus the only linear combination giving the zero vector is the trivial linear combination, which proves that the vectors are linearly independent.

8.7.6 This follows at once from the CBS Inequality by putting

$$\vec{v} = \begin{bmatrix} \frac{a_1}{1} \\ \frac{a_2}{2} \\ \cdots \\ \frac{a_n}{n} \end{bmatrix}, \quad \vec{u} = \begin{bmatrix} 1 \\ 2 \\ \cdots \\ n \end{bmatrix}$$

and noticing that

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}.$$

Bibliography

[Cul] CULLEN, C., **Matrices and Linear Transformations**, 2nd ed., New York: Dover Publications, 1990.

[Del] DELODE, C., **Géométrie Affine et Euclidienne**, Paris: Dunod, 2000.

[Fad] FADDÉEV, D., SOMINSKY, I., **Recueil d'Exercises d'Algèbre Supérieure**, 7th ed., Paris: Ellipses, 1994.

[Hau] HAUSNER, M., **A Vector Space Approach to Geometry**, New York: Dover, 1998.

[Lan] LANG, S., **Introduction to Linear Algebra**, 2nd ed., New York: Springer-Verlag, 1986.

[Pro] PROSKURYAKOV, I. V., **Problems in Linear Algebra**, Moscow: Mir Publishers, 1978.

[Riv] RIVAUD, J., **Ejercicios de álgebra**, Madrid: Aguilar, 1968.

[Tru] TRUFFAULT, B., **Géométrie Élémentaire: Cours et exercises**, Paris: Ellipses, 2001.

GNU Free Documentation License

Version 1.2, November 2002
Copyright © 2000,2001,2002 Free Software Foundation, Inc.

51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA

Everyone is permitted to copy and distribute verbatim copies of this license document, but changing it is not allowed.

Preamble

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "**Document**", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "**you**". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "**Modified Version**" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "**Secondary Section**" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "**Invariant Sections**" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "**Cover Texts**" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "**Transparent**" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "**Opaque**".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "**Title Page**" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "**Entitled XYZ**" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "**Acknowledgements**", "**Dedications**", "**Endorsements**", or "**History**".) To "**Preserve the Title**" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties–for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

Que a quien robe este libro, o lo tome prestado y no lo devuelva, se le convierta en una serpiente en las manos y lo venza. Que sea golpeado por la parálisis y todos sus miembros arruinados. Que languidezca de dolor gritando por piedad, y que no haya coto a su agonía hasta la última disolución. Que las polillas roan sus entrañas y, cuando llegue al final de su castigo, que arda en las llamas del Infierno para siempre.

-Maldición anónima contra los ladrones de libros en el monasterio de San Pedro, Barcelona.